UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Free-moving Omnidirectional 3D Gamma-ray Imaging and Localization

Permalink
https://escholarship.org/uc/item/39f269gk

Author
Hellfeld, Daniel

Publication Date
2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Free-moving Omnidirectional 3D Gamma-ray Imaging and Localization

by

Daniel Hellfeld

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Nuclear Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kai Vetter, Chair
Professor Jasmina Vujic
Professor Steven Conolly
Adjunct Professor Lee Bernstein

Summer 2019

Free-moving Omnidirectional 3D Gamma-ray Imaging and Localization

Copyright 2019
by
Daniel Hellfeld

Abstract

Free-moving Omnidirectional 3D Gamma-ray Imaging and Localization

by

Daniel Hellfeld

Doctor of Philosophy in Nuclear Engineering

University of California, Berkeley

Professor Kai Vetter, Chair

The ability to localize and map the distribution of gamma-ray emitting radionuclides in 3D has applications in medical imaging, nuclear contamination remediation, and nuclear security and safeguards. The deployment of freely moving detection systems, such as hand-held instruments or ground/aerial-based vehicles, are critical in overcoming the inverse square law and complex shielding scenarios. Using auxiliary contextually-aware sensors, capable of perceiving spatiotemporal characteristics of the environment, these systems can simultaneously generate 3D maps of the surroundings and track the position and orientation of the gamma-ray sensitive detectors in the scene. The fusion of contextual scene data and gamma-ray detector data to facilitate real-time 3D gamma-ray image reconstruction has previously been demonstrated with mobile germanium and CdZnTe-based Compton cameras for gamma-ray energies ranging from a few hundred keV to several MeV. This concept is applied here for lower energy ($50-400$ keV) gamma-rays using an active coded mask imaging modality. The platform for demonstration is the Portable Radiation Imaging Spectroscopy and Mapping (PRISM) system, which is a hand-held spherical active coded array of many 1 cm^3 coplanar-grid CdZnTe detectors designed for omnidirectional coded mask and Compton imaging and uniform directional sensitivity. This work presents the design, development, and coded mask optimization of PRISM, as well as the methodologies developed for real-time reconstruction using a scene data constrained, GPU-accelerated, list-mode maximum likelihood expectation maximization (ML-EM) algorithm. Experimental results from several measurements in the lab and in the field are shown.

A novel approach to 3D gamma-ray image reconstruction for scenarios where sparsity in the source distribution may be assumed, for example radiological source search, is also presented. While the generality of ML-EM enables use in a wide variety of scenarios, it is susceptible to overfitting, limited by the discretization of spatial coordinates, and can be computationally expensive. A more well-conditioned Point-Source Localization (PSL) approach is formulated as an optimization problem where both position and

source intensity are continuous variables. This formulation is then extended and generalized to an iterative algorithm for sparse parametric 3D image reconstruction called Additive Point-Source Localization (APSL), where the image is considered the sum of multiple point-sources whose position and intensity are continuous in nature. APSL mitigates overfitting in its iterative bottom-up nature and statistically-founded stopping criteria and, because of the inherent point-source assumption and continuous variables, results in images with improved accuracy and interpretability as compared with ML-EM. A set of simulated source search scenarios using a single non-directional detector is considered to demonstrate the concept and compare ML-EM and APSL. Experimental results using a nearly isotropic, contextually-aware, $LaBr_3$ detector system are then presented, finding improved localization accuracy and computational efficiency with APSL.

To Kalie and Eleanor.

Contents

List of Figures

List of Tables

List of Abbreviations

ADC Analog-to-Digital Converter

AIC Akaike Information Criterion

APSL Additive Point-Source Localization

ASIC Application Specific Integrated Circuit

BIC Bayesian Information Criterion

BOBYQA Bound Optimization BY Quadratic Approximation

CAPTAN Compact And Programmable daTa Acquisition Node

CDF Cumulative Distribution Function

CLLBC $Cs_2LiLaBr_{6-x}Cl_x$:Ce

CLYC Cs_2LiYCl_6:Ce

CPG Coplanar Grid

CPU Central Processing Unit

CZT $Cd_{1-x}Zn_xTe|_{x=0.1}$

DOI Depth-of-Interaction

FBP Filtered Back-Projection

FF-FD Far-Field Finite-Detector

FOV Field-of-View

FPGA Field Programmable Gate Array

FWHM Full-Width-at-Half-Maximum

GD Great Deluge

GPS Geopositioning System

GPU Graphics Processing Unit

HEALPix Hierarchical Equal Area isoLatitude Pixelation of a sphere

HEMI High Efficiency Multimodal Imager

HPGe High-Purity Germanium

IMU Inertial Measurement Unit

INS Inertial Navigation System

LBNL Lawrence Berkeley National Laboratory

LiDAR Light Detection and Ranging

MAP Maximum a Posteriori

ML-EM Maximum Likelihood Expectation Maximization

MURA Modified Uniformly Redundant Arrays

NF Near-Field

NLOPT Non-Linear Optimization Library

NRA Non-Redundant Arrays

OS-EM Ordered Subsets Expectation Maximization

OSL One Step Late

PNP Psuedo-Noise Product Array

PRISM Portable Radiation Imaging Spectroscopy and Mapping

PSD Pulse Shape Discrimination

PSF Point Spread Function

PSL Point-Source Localization

PYGMO PYthon parallel Global Multiobjective Optimizer

RAM Random Access Memory

RGB-D Red-Green-Blue-Depth

ROI Region of Interest

SA Simulated Annealing

SBP Simple Back-Projection

SBR Signal-to-Blur Ratio

SDF Scene Data Fusion

SLAM Simultaneous Localization and Mapping

SNR Signal-to-Noise Ratio

SPECT Single Photon Emission Computed Tomography

TV Total Variations

UAS Unmanned Aerial System

URA Uniformly Redundant Arrays

Acknowledgments

I would like to thank many people who helped make this dissertation possible. First, my advisor Kai Vetter, for creating such a unique and productive research environment in the Applied Nuclear Physics (ANP) program at Berkeley Lab and the Berkeley Applied Research on the Imaging of Neutrons and Gamma-rays (BeARING) group at UC Berkeley. Throughout this work, I have been able to be involved in the development of multiple detectors systems spanning multiple projects. This has given me the opportunity to work with and learn from almost everyone in ANP in one way or another, providing me with valuable insight as to how different projects are proposed, managed and executed. I have also been able to be involved in and give lectures for several undergraduate courses, which I found to be both fun and rewarding. Finally, I was fortunate enough to travel to Fukushima, Japan, and participate in several workshops relating to the 2011 accident. None of this would have been possible without Kai's support and encouragement.

Paul Barton, for his excellent mentorship both in and out of the lab over the years, including teaching me how to solder. Lucian Mihailescu, for sparking new ideas and for always being available to discuss any aspect of imaging. Don Gunter, for his mathematical expertise that has proven useful in many aspects of this work. Andy Haefner, for mentoring me in all things imaging and for always guiding me to answer my own questions. Tenzing Joshi and Mark Bandstra, for involving me and working with me in the development of new ideas.

The remaining members of the PRISM team including Ross Barnowski, Victor Negut, Marcos Turqueti and Alex Moran, for making the system a reality. Ryan Pavlovsky, Joey Curtis, Erika Suzuki, Ren Cooper, Brian Quiter and the rest of ANP, for always being a resource to me. My cube-mate and espresso partner, Kyle Bilton, for always being a source of interesting ideas, especially over clam chowder at our scheduled lunch meetings. Joe Vanderlip, for assisting me in experimental data collection and motivating me to get out of the office and exercise. My doctoral committee - Jasmina Vujic, Lee Bernstein and Steven Conolly - for agreeing to be a part of and guiding me through this process. The Nuclear Science and Security Consortium, for supporting my entire graduate career, including travel to domestic and international conferences where I have made long-lasting connections that will undoubtably serve me well as I continue my career in this field. The Defense Threat Reduction Agency, specifically Hank Zhu, for funding the PRISM project and for providing opportunities to demonstrate the system in the field and receive feedback from potential end-users.

I want to acknowledge and express my sincere appreciation to my parents, Hank and Eileen, for their support through all (10 years) of my higher education. Finally, I want to thank my wife, Kalie, for her love and patience.

This material is based upon work supported by the Department of Energy National Nuclear Security Administration under Award Number DE-NA0003180 and by the Defense Threat Reduction Agency under Contract Numbers 10027-28018 and 10027-19346.

This Ph.D. thesis was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Chapter 1

Introduction

This dissertation is centered around three major achievements in radiation detection instrumentation and gamma-ray imaging: the design and optimization of an omnidirectional spherical active coded mask gamma-ray imager, the implementation and demonstration of free-moving 3D active coded mask imaging with scene data fusion, and the development of a novel approach to general sparse parametric 3D image reconstruction using Poisson likelihood. The following three sections provide brief introductions into these topics, with more detailed discussions provided in later chapters.

1.1 Gamma-ray Detection, Localization and Imaging

The detection, identification, characterization, localization and mapping of gamma-ray source distributions have broad applications in numerous fields including, but not limited to, gamma-ray astronomy [1]–[3], medical imaging [4]–[7], contamination remediation [8], [9], and nuclear security, safeguards and nonproliferation [10], [11]. Coded aperture imaging, originally developed by [12]–[14], has been widely used in these areas for X-ray and gamma-ray imaging in the energy regime of tens of keV to a few hundred keV [15]. This low-energy regime is critical in nuclear security and safeguards applications (the primary focus of this work) as many radionuclides of interest emit low-energy (< 450 keV) gamma-rays. For example, ^{241}Am, a byproduct of the buildup of ^{239}Pu in nuclear reactors, and ^{235}U emit gamma-rays at 60 and 186 keV, respectively.

Coded aperture is a mechanical collimation imaging technique that maintains the high angular resolution of a pinhole imager while increasing the image signal-to-noise ratio (SNR) [16]. Typical systems utilize a passive high-Z (e.g., Pb, W) mask in front of a position-sensitive gamma-ray detector. The mask is designed to uniquely modulate the incident gamma-ray flux based on the source direction (i.e., the spatial pattern of events in the detector uniquely determines the direction of the source). The reconstruction of the estimated source distribution is typically limited to a region-of-interest (ROI) in

energy, using only full-energy absorption (i.e., photopeak) events. The effectiveness of the coded mask therefore decreases with increasing energy as the attenuation of photons weakens and the detection pattern blurs.

The replacement of passive mask materials with active elements has been explored [17]–[19] to increase detection sensitivity, decrease weight, and enable Compton imaging of higher energy gamma-rays. The typical planar design of these systems, however, is prone to imaging artifacts, has a non-uniform Compton efficiency and lever-arm distribution, and the coded aperture modality suffers from a limited coded field-of-view (FOV), resulting in "point-and-shoot" type operation. This work proposes a novel imaging design by rearranging the detectors into an active coded spherical configuration, resulting in an omnidirectional FOV and isotropic efficiency for both coded aperture and Compton imaging. The Portable Radiation Imaging Spectroscopy and Mapping (PRISM) system, designed and assembled entirely at Lawrence Berkeley National Laboratory (LBNL), is a hand-held, omnidirectional, dual-mode (coded and Compton), spherical active coded array ($\varnothing$14 cm) of 1 cm^3 room-temperature operated coplanar grid (CPG) $Cd_{1-x}Zn_xTe|_{x=0.1}$ (CZT) detectors. Part of this work describes the design and development of two prototype PRISM systems. This includes the optimization of the number and configuration of detectors for 4π active coded mask imaging and imaging improvements with depth-of-interaction (DOI) read-outs. While the system is capable of Compton imaging, this work focuses primarily on the active coded mask imaging modality.

1.2 Free-moving 3D Imaging and Scene Data Fusion

Free-moving systems are needed to efficiently localize sources in potentially cluttered environments or complex shielding scenarios as well as to quickly map radiation fields from small scales (e.g., rooms, buildings) to large scales (e.g., streets, blocks, cities). The distinction of a free-moving system is made here to discriminate between portable systems that are designed to acquire data in several static positions and mobile systems designed to acquire gamma-ray data while the system is in motion, without any restrictions on the position or orientation of the detector.

In radiological source search, hand-held free-moving systems are important in overcoming the inverse square law to get much closer to weak sources or to view suspected source locations from various perspectives in order to circumvent potential shielding (e.g., finding narrow streaming paths) [20]–[23]. In nuclear contamination remediation and consequence management, unmanned ground-based or aerial-based systems facilitate quick wide-area mapping of gamma-ray source distributions in potentially difficult-to-navigate environments, while limiting dose to human operators [24]–[28]. These maps can be used as a guide for further remediation, to ensure the area is no longer contami-

nated, or for contamination avoidance planning.

Traditionally, such mobile systems consist of non-imaging detectors and produce proximity images or maps, using only the gamma-ray flux modulation with distance as a means for localization or source distribution estimation [28], [29]. Imaging capabilities, such as coded aperture and Compton imaging, can significantly enhance the localization accuracy of sources and improve the mapping of complex distributions with higher image contrast. However, current state-of-the-art portable gamma-ray imaging systems [30], [31] are designed for static, long dwell, data acquisition and produce 2D images. Three-dimensional imaging can be achieved simply by combining data from several measurements taken at different locations around the source. This is standard practice in medical imaging, where a detector system is rotated around a stationary, voxelized image space. The gantry-mounted detector follows a fixed and known trajectory, therefore the system position and orientation (i.e., pose) is known precisely at every measurement. The data collected at each measurement can be then appropriately translated and rotated into the global coordinate system of the static image space and combined for 3D image reconstruction. Free-moving systems, however, do not follow well-characterized trajectories around a pre-defined image space and therefore must rely on other methods for pose estimation and adaptive image voxelization in order to perform 3D imaging.

Mihailescu et al. [32] first demonstrated 3D imaging with a mobile system by acquiring data from multiple static locations around a source and manually determining the system pose at each measurement using a laser rangefinder. Similarly, Kishimoto et al. [33] performed measurements from four known static positions around a fixed imaging space. More recent work has developed the use of simultaneous localization and mapping (SLAM) [34], [35] for real-time automated pose tracking of continuously moving mobile systems. An excellent discussion on the development and use of SLAM in this context can be found in [36].

In addition to pose estimation, SLAM generates a 3D model of the environment around the detector. The 3D scene model can be used to define and provide context to the 3D gamma-ray image space as well as to constrain the image reconstruction (to surfaces in the 3D scene), increasing image accuracy, decreasing noise, and improving the computational efficiency of the reconstruction. This approach is referred to as scene data fusion (SDF) [36].

The contextual tracking and mapping as well as gamma-ray imaging has been demonstrated to perform in real-time, providing a fully 3D scene and gamma-ray model to the user during a measurement [19]. This information allows for course correction, decreasing the overall time to detect and image suspected sources. The real-time capability can also facilitate autonomous search with unmanned systems using the gamma-ray image as input to a route optimization routine [37].

The work presented here is built upon previous work [19], [32], [36] which focused

on the Compton imaging modality for gamma-ray energies from several hundred keV to a few MeV. One of the primary objectives of this work is to apply the 3D SDF approach to the active coded mask modality for lower gamma-ray energies (tens of keV to a few hundred keV), using PRISM as the detector platform for demonstration. Real-time 3D imaging is achieved using the maximum likelihood expectation maximization (ML-EM) reconstruction algorithm [38] in the list-mode formalism [39] with graphics processing unit (GPU) acceleration. Several reconstruction scenarios are considered, both in the lab and in the field, with source activities ranging from μCi to mCi, gamma-ray energies from 60 to 356 keV, and source positions in the near- and far-field regimes.

1.3 Sparse Parametric Image Reconstruction

Often in the case of radiological source search, the task is to localize and quantify the activity of one or more compact sources of gamma-ray radiation in an unknown environment. In the maximally general case (unknown background and no priors on the source distribution), this problem can be solved with the methods briefly described above (i.e., discretizing the spatial dimensions of the image space and employing the ML-EM algorithm). While the generality of the traditional grid-based ML-EM formulation enables its use in a wide variety of scenarios, there are two significant disadvantages. First, the images are limited in accuracy by the degree of spatial discretization (voxel sizes typically range from $10-30$ cm [19], [40]). Finer discretization can be used, but this quickly drives the reconstruction to become computationally expensive and memory intensive. Second, and more importantly, the sparse reconstruction problem presented to ML-EM is generally ill-posed (i.e., the number of gamma-ray intensity estimates is much larger than the number of measurements) and thus is susceptible to overfitting. While this work demonstrates that well-conditioned reconstructions can be achieved through the use of highly directional imaging systems and SDF regularization, these sophisticated systems tend to be expensive. The additional cost can preclude their use in some cases, favoring simpler, not necessarily directional, free-moving systems. A novel approach to 3D gamma-ray image reconstruction is proposed and developed here with this case in mind, demonstrating that advanced gamma-ray imaging systems may not in fact be necessary for the sparse search scenario.

The approach first reformulates traditional ML-EM with a single point-source assumption, where both the position and source intensity are treated as continuous variables. This formulation is then extended to an iterative algorithm called Additive Point-Source Localization (APSL) for general sparse parametric 3D image reconstruction, where the image is considered the sum of multiple point-sources. APSL mitigates overfitting in its iterative bottom-up nature and statistically-founded stopping criteria and produces more accurate images with a reduced computational burden compared to ML-EM. A set

of simulated and experimental measurements are considered for comparison.

1.4 Dissertation Structure

This dissertation is organized as follows:

- Chapter 2 covers introductory concepts including gamma-ray detection and imaging principles, image reconstruction algorithms and acceleration techniques, contextual sensing and SDF.
- Chapter 3 describes the PRISM system including prototype design, development, and mask optimization.
- Chapter 4 shows results for omnidirectional 2D and real-time 3D imaging with PRISM in the far-field and near-field and imaging improvements with DOI.
- Chapter 5 presents the development and demonstration of the APSL approach for sparse parametric 3D gamma-ray imaging.
- Chapter 6 includes general conclusions and future outlook of this work.

1.5 Relevant Publications

Text and figures from the following papers, of which I was the primary author, are included in this dissertation with the permission of all authors:

- **D. Hellfeld**, P. Barton, D. Gunter, L. Mihailescu, and K. Vetter, "A Spherical Active Coded Aperture for 4π Gamma-Ray Imaging", IEEE Transactions on Nuclear Science, vol. 64, no. 11, pp. 2837-2842, 2017. [41]
- **D. Hellfeld**, T. H. Y. Joshi, M. S. Bandstra, R. J. Cooper, B. J. Quiter, and K. Vetter, "Gamma-Ray Point-Source Localization and Sparse Image Reconstruction using Poisson Likelihood", IEEE Transactions on Nuclear Science (in press), 2019. [42]
- **D. Hellfeld**, P. Barton, D. Gunter, A. Haefner, L. Mihailescu, and K. Vetter, "Real-time Free-moving Active Coded Mask 3D Gamma-ray Imaging", Submitted to IEEE Transactions on Nuclear Science (under review), 2019. [43]

Chapter 2

Concepts

2.1 Gamma-ray Sources and Interactions

Before gamma-ray imaging can be discussed, it is important to understand the mechanisms in which gamma-rays are created and interact in matter. This section briefly introduces these concepts, within the context of gamma-ray imaging in the energy range of tens of keV to hundreds of keV. Most of this section is referenced from [44], in which more detailed discussions of these concepts can be found. Additional references on this subject include [45]–[48].

Gamma-ray radiation is emitted in the process of nuclear de-excitation, or the transition of nuclei from excited states to lower-lying nuclear levels. Gamma-rays are therefore distinct from X-rays only in the origin of production (atomic vs. nuclear transitions). Nuclei can be left in excited states following the decay of a parent radionuclide (e.g., β-decay) or the results of a nuclear reaction. Gamma-rays produced from β-decay are typically limited to energies below 2.8 MeV while higher energies can be produced with specific nuclear reactions. Nuclear states have well-defined energies, therefore gamma-ray energies are also well-defined and shed light on the nuclear structure of the exited nucleus. While the lifetimes of these excited states are small (typically on the order of picoseconds or less), the natural line-widths of the energy distribution are always less than the energy resolution of spectroscopic detectors. However, significant spread of the distribution can occur in the de-excitation of nuclei with large velocities (i.e., Doppler broadening). This effect is specific to certain processes and reactions and is typically not of concern for the cases presented in this work, therefore any observed broadening of the photon energy distribution is attributed to the inherent resolution of the detector due to statistical uncertainties associated with the creation of finite information carriers.

Once produced, gamma-rays interact with matter primarily through four distinct processes that result in discrete transformations that can include a change in energy, direction, or particle type. Figure 2.1 displays the individual interaction mass atten-

uation coefficients, μ/ρ (units of cm^2/g) as a function of energy, shown here specifically in the CZT detector material. Interaction cross-sections can be computed by $\sigma = (\mu/\rho)(m/N_A)$, where m is the molar mass and N_A is Avogadro's number. Each interaction process will be discussed individually.

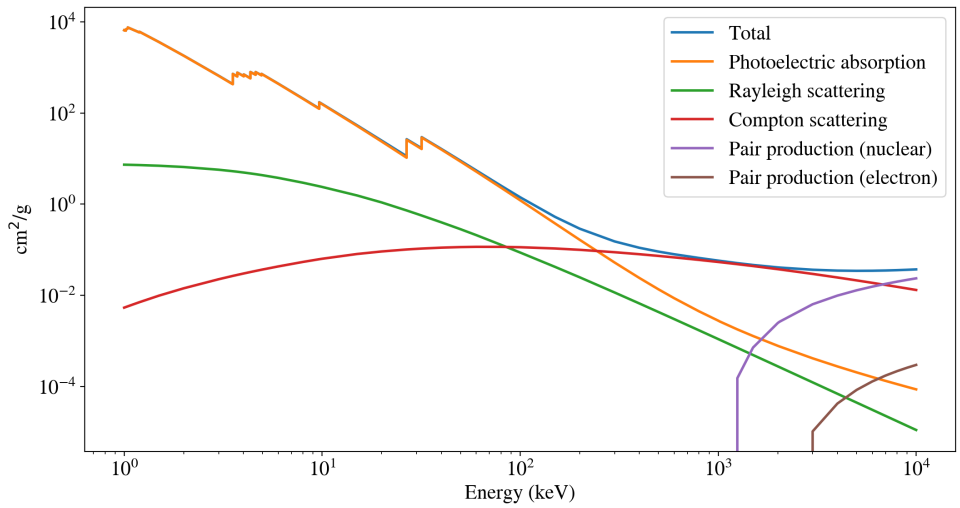

Figure 2.1. Gamma-ray interaction mass attenuation coefficients (units of cm^2/g) in CZT.

Photoelectric absorption is the interaction of a photon with an atom in which the incident energy of the photon is completely absorbed and a photoelectron is ejected from one of the bound electron shells, with kinetic energy equal to the photon energy less the binding energy of the shell from which it was ejected. The resultant vacancy in the electron shell is quickly filled either from a nearby free electron or from an outer-shell electron, emitting one or more characteristic X-rays or Auger electrons with energy equal to the difference in binding energies. Typically, the kinetic energies of the photoelectron and additional particles are deposited nearby the original interaction site and the entire energy of the incident photon is accounted for. In some cases, the characteristic X-ray(s) can escape from the detecting medium, resulting in an energy deposition less than the total incident photon energy. The nucleus will also experience a recoil due to the conservation of momentum, however this effect is negligible for the interaction energies dominated by photoelectric absorption. The cross-section is highest for low-energy photons and in materials with high atomic number (Z). This interaction process is critical to coded aperture and active coded mask imaging as, given a known source energy, a full energy deposition in the detector, in most cases, indicates that the photon direction was unchanged from the initial emission.

Compton scattering is the interaction of a photon with an electron (assumed free and at rest) in which the photon experiences a change in direction and a subsequent energy

loss (imparted to the recoil electron). The energy loss ($E_\gamma - E'_\gamma$, where E_γ and E'_γ are the energies of the incident and scattered photons, respectively) is coupled to the scattering angle θ between the incident and scattered photon and can range from zero (in forward scattering, $\theta = 0$) to a significant fraction of the incident energy (in back scattering, $\theta = \pi$). The scattered gamma-ray energy is given by

$$E'_\gamma = \frac{E_\gamma}{1 + \frac{E_\gamma}{m_e c^2}(1 - \cos\theta)}, \tag{2.1}$$

where $m_e c^2$ is the rest mass energy of the electron (511 keV). If the assumption of a free electron at rest is removed, the energy of the scattering photon is no longer unique to the scattering angle and will instead follow a distribution of energies with finite width about that energy. The cross-section increases linearly with Z (scaling with the number of available electrons in the medium) and gradually increases with increasing energy to a maximum that depends on the material (e.g., few hundred keV in CZT) followed by a gradual decrease at higher energies. Compton scattering represents a noise channel to coded mask imaging as photons can scatter before photoelectric absorption, obfuscating either just the energy (source photon scatters in detector and escapes) or both the energy and direction (source photon scatters before entering the detector or multiple events occur in the detector) of the source. Gamma-ray tracking, or the ability to detect, discriminate, and attribute multiple events in a detector system, can be used to reject Compton scattering events from mechanical collimation or photoelectric based imaging methods or to reconstruct possible source directions (i.e., Compton imaging).

In addition to Compton scattering, photons can also change directions through coherent scattering, also referred to as Rayleigh scattering. In this process, the photon interacts with the entire atom without excitation or ionization, retaining all the incident energy. The cross-section is highest at lower energies and in high-Z materials. While this represents another noise channel for coded mask imaging, the cross-section is typically $< 10\%$ of the total interaction cross-section and in most cases can be ignored.

The final interaction discussed here is pair production. If the photon energy exceeds twice that of the rest mass energy of an electron (i.e., 1022 keV), it is energetically possible, in the presence of the Coulomb field of the nucleus[1], for the photon to transform into an electron-positron pair with any excess energy imparted as kinetic energy shared between the pair. The positron will eventually deposit all of its kinetic energy and subsequently annihilate with an electron in the medium, producing two annihilation photons. The two annihilation photons typically have energies of 511 keV and are emitted back-to-back (180°) due to conservation of momentum, however, these characteristics

[1]At high enough energies, it is also possible for the photon to interact with the atomic electron field and pair produce. Since a significant amount of momentum is imparted to the interacting electron, this process is typically called triplet production [49].

may fluctuate depending on the existence of finite momentum in the annihilation and if the annihilating electron is bound. The pair-production cross-section increases approximately by the square of Z and sharply with increasing energy, though does not dominate the total cross-section until very high energies (several MeV). As coded mask imaging is designed for low-energy gamma-rays, this interaction is generally ignored and will not be discussed here any further.

2.2 Detection Principles

Imaging relies first and foremost on the detection of gamma-rays. In the context of this work, detection refers to the measurement of gamma-ray interaction parameters including the position, time, and deposited energy of the interaction in a medium. While the process of extracting this information varies with material (i.e., gas, scintillator, semiconductor, etc.), detection is fundamentally based on the production, transport, and collection of information carriers. The information carriers ultimately produce an electrical signal that can be read out and analyzed. The design of gamma-ray detectors (i.e., the composition and configuration of the interaction medium as well as the accompanying read-out electronics) define the accuracy and precision in the measurement of the parameters of interest, limited by the physics and subsequent statistics of the information carriers.

This work focuses primarily on the use of the semiconductor material CZT, therefore a brief introduction into gamma-ray detection in the context of CZT is provided here. Detailed information on semiconductor detectors can be found in [50] and further discussion of other detector materials can be found in [44]–[48].

The secondary charged particle resultant from an ionizing gamma-ray interaction (e.g., the electron) in a semiconductor will produce information carriers (electron-hole pairs) as the particle deposits energy along its track. Semiconductor material is characterized by the small, $\mathcal{O}(\text{eV})$, energy required to liberate an electron from the valence band to the conduction band. This small band-gap facilitates the production of a large number of liberated electrons in the conduction band (and holes in the valence band) as the secondary charged particle gradually slows and deposits energy in the medium. High voltage is applied across electrical contacts (i.e., electrodes), creating a large electric field that drifts the electrons and holes to the anode and cathode electrodes, respectively. The drift of the charge carriers creates an measurable current that can be shaped and amplified with read-out electronics. The integrated current is a measure of the total induced charge on the electrodes, proportional to the number of initially created charge carriers and thus energy deposition. Factors that can degrade this proportionality include charge trapping and recombination (from crystal defects and impurities) and charge diffusion. The large number of electron-hole pairs is what separates semiconductor detectors from

other spectroscopic detection media, as the statistics facilitate high precision in the measurement of total energy deposition. Other advantages of semiconductors include the high density and Z leading to highly efficient detectors, and the small charge transit time leading to fast timing characteristics. Some disadvantages include the small size due to limitations in the crystal growth process and the susceptibility to radiation damage.

The *gold-standard* of spectroscopic gamma-ray detectors with efficiencies large enough for nuclear security applications is high-purity Germanium (HPGe). Germanium is a dense high-Z semiconductor material (band-gap of $\sim$0.7 eV) capable of efficient gamma-ray detection with superior energy resolution, nominally $<$ 0.2% full-width-at-half-maximum (FWHM) at 662 keV. Advances in zone refining and crystal growth methods, starting in the 1970's, eventually lead to the production and ultimate widespread adoption of HPGe detectors with impurity levels down to 10^9 atoms per cm^3 and large sensitive volumes relevant to applications in electron and gamma-ray spectroscopy. While dependent on the detector geometry and applied bias, the high charge mobility (speed of drift) and lifetime (average time before trapping or recombination) of both charge carriers facilitate fast timing characteristics for full charge collection. Position sensitivity is also possible in HPGe detectors through different electrode schemes, with orthogonal strips being a common example [2], [36], [51], [52]. While the properties of HPGe present several advantages for gamma-ray spectroscopy, the small band-gap requires the detector to be kept cold (typically liquid nitrogen temperatures at 77 K) to reduce thermal excitation of electrons to the conduction band and leakage current. Operationally this requires an accompanying complex cryostat unit with either a liquid nitrogen dewar or mechanical/electrical cooler, resulting in cumbersome detector systems. The temperature requirements can be mitigated by utilizing materials with larger band-gaps, however at the cost of energy resolution.

A common alternative semiconductor material to germanium is CZT. The larger band-gap energy ($\sim$1.5 eV) translates to worse energy resolution performance (nominally $<$ 2% at 662 keV, though $<$ 1% can be achieved depending on the electrode scheme), however it facilitates room-temperature operation and removes the need for cooling equipment that, in addition to weight, can require large amounts of power and present significant attenuating and scattering material around the gamma-ray sensitive volume. The mobility and lifetime of both charge carriers is also much worse in CZT than in HPGe, forcing smaller crystal sizes to achieve reasonable timing properties. The most significant disadvantage of CZT is the poor hole collection properties. The holes cannot drift across the crystal over any reasonable timescale for gamma-ray detection and results in an overall incomplete charge collection. The holes effectively remain within the bulk of the crystal and induce charge on the anode (degrading the signal from the electrons). Moreover, this effect is depth-dependent, meaning that, for a given energy deposition, the induced charge on the cathode and the amount of required charge compensation on the anode will vary with interaction position along the anode-cathode axis.

This problem has been solved with various electrode schemes, most of which utilize unipolar charge sensing techniques.

One electrode scheme used to mitigate the problems associated with poor hole mobility in CZT is the CPG [53]–[56]. This is comprised of a typical planar cathode and an interdigitated anode grid with a collecting grid and non-collecting grid. A voltage is applied (nominally -80 V) between the two grids to force the electrons towards the collecting grid electrode. The holes remain effectively static during the electron collection time and induce the same depth-dependent charge on both grids. To first order, the induced charge is assumed to vary linearly with the distance to the anode. The signal difference between the collecting and non-collecting grids removes the induced charge from the holes, but remains proportional to the number of initially generated charge carriers and is derived only from the electrons. However, the electrons will also suffer from a depth-dependent charge loss due to trapping. A linear differential gain adjustment is made on the non-collecting grid to effectively reintroduce some of the depth-dependent charge from the holes. Assuming no significant spatial non-uniformities in the crystal, the overall induced signal from the electrons is no longer a function of the interaction depth though the overall induced charge will be less than the total charge originally produced [18].

In practice, the trapping is an exponential function with distance and thus the linear differential gain method only works for small detector thicknesses (less than the electron drift length). In this case, event-by-event depth estimates are required to correct for non-linear behavior [57]. The holes will also induce a depth-dependent charge on the cathode. If read out, the ratio of the cathode signal to the depth-independent anode grid signal provides an estimate of the DOI. Note this is only a 1D position sensitivity.

While the CPG technique is utilized for the CZT detectors used in this work, another common approach to unipolar charge sensing in CZT uses a pixellated anode [58] and the *small pixel effect*, in which the induced charge on the anode pixels occur only when the charges are spatially close to the pixels. This technique has been shown to be capable of superior energy resolution to CPG [59] and can facilitate event-by-event 3D position sensitivity [60]. However, these detectors come at the cost of increased number and complexity of read-out electronics and can therefore be cumbersome in some applications.

2.3 Imaging Principles

The detection of gamma-rays does not necessarily facilitate the ability to determine the direction of the incident photon. The term imaging is used here as the process of determining the origin, defined as either a position in $\mathbb{S}^2$ (i.e., a direction) or in $\mathbb{R}^3$ (i.e., an XYZ location), of detected gamma-rays. In contrast to visual photography, gamma-

rays do not experience any significant amount of refraction and thus typically rely on lens-less imaging techniques[2]. This is accomplished through collimation, either mechanical collimation (i.e., design a configuration of passive and active elements to physically modulate the gamma-ray signal in some particular manner) or kinematical collimation (i.e., selecting events that satisfy some kinematic interaction criteria in the detector and using physics to constrain the incident source direction).

2.3.1 Passive Coded Aperture

The simplest mechanical collimation technique is a pinhole camera. In this approach, a large piece of high-Z (e.g., Pb, W) gamma-ray attenuating material, or a *mask*, is placed at some distance in front of a position sensitive detector. A single small opening is introduced into the mask to create an *aperture*. Assuming no diffraction (100 keV = 12 pm), the pinhole aperture creates a unique one-to-one mapping between the interaction position in the detector and the incident photon direction. In this simple treatment, a 1D camera would be able to reconstruct the direction of a source perfectly (i.e., the point spread function (PSF) is a delta function). The significant attenuation drastically limits the signal and thus long exposure times are required to produce images with appreciable SNR. Therefore this approach is not suitable for weak sources or large standoffs where statistics are limited. The pinhole can be enlarged or multiple pinholes can be introduced into the mask to increase the SNR, but the unique mapping will degrade and degeneracies will be introduced into the solution (i.e., the PSF blurs and may contain artifacts). However, if multiple pinholes are arranged into a specific pattern, then an ensemble of events will create a detection pattern, or *shadowgram*, that is unique to source location. The use of a multi-hole pattern, or *coded aperture*, was originally developed by [12]–[14] and has been widely used in areas such as X-ray and gamma-ray astronomy, medical imaging, and nuclear security for radiation imaging [1], [15]. The technique maintains the high angular resolution of a pinhole imager while increasing the image SNR [16] proportional to the number of holes in the mask.

As in a vast variety of physical processes, the imaging system can be described by a convolution. In order to avoid working with reflected images, it is useful to instead describe the system as a correlation. In one dimension, the correlation ($\star$) of two functions is given by

$$f(x) \star g(x) = \int_{-\infty}^{\infty} du\, g^{\dagger}(u-x) f(u)\,, \tag{2.2}$$

where $g^{\dagger}(x)$ is the complex conjugate of $g(x)$. Correlation can be related to the convolution operation ($*$) by the adjoint operation (or simply a reflection operation, as all

[2]Gamma-ray focusing technologies such as multilayer-coated mirrors [61] and Laue lenses [62] have been developed for space-borne nuclear astrophysics measurements. These techniques are outside the scope and scale of the terrestrial measurements in this work.

functions are strictly real in this case)

$$f(x) * g(x) = \int_{-\infty}^{\infty} du\, g(x-u) f(u)\,. \tag{2.3}$$

The imaging equation can then be written as

$$P = A \star O + N\,, \tag{2.4}$$

where P is the detected image, A is the aperture, O is the source object, and N is noise [16]. Using the correlation theorem (similar to the convolution theorem), the source object can be solved for analytically

$$\hat{O} = R\mathcal{F}^{-1}\left[\frac{\mathcal{F}(P)}{\mathcal{F}(A)}\right] = O + R\mathcal{F}^{-1}\left[\frac{\mathcal{F}(N)}{\mathcal{F}(A)}\right]\,, \tag{2.5}$$

where $\hat{O}$ is the reconstructed source object, R is the reflection operator, $\mathcal{F}$ is the Fourier Transform, and $\mathcal{F}^{-1}$ is the inverse Fourier Transform. This method of *deconvolution* is simple and straightforward. However, since the aperture is typically represented as a binary array of ones and zeros, the $\mathcal{F}(A)$ term can be small, leading to large noise gain in the image. The cross-correlation method [16] instead uses the mask pattern to define a post-processing array G such that $A \star G$ approximates a delta function. The post-processing array is applied to Eq. 2.4 to solve for O. The noise term is still present, however it will not contain singularities as in the deconvolution method. While $A \star G$ is generally not a delta function for random apertures (and thus introducing image blur), the uniformly redundant array (URA) and the balanced correlation method to solve for G introduced by [16] have been shown to effectively create a scenario in which $A \star G \approx \delta$.

2.3.2 Active Coded Mask

The passive aperture approach presents several problems for imaging systems of relevance for radiological source search and distributed source mapping. First, the reconstruction relies on a significant fraction of gamma-ray events being discarded (URAs utilize a 50% open fraction). This is not ideal in count-starved scenarios such as weak or shielded source search, where statistics are important to solve the detection problem. Second, the weight of the mask material can result in cumbersome detectors, limiting the overall mobility of free-moving systems. The thickness of the mask typically limits the imageable energies to tens of keV to a few hundred keV. Third, the traditional planar construction of these types of imagers suffers from a severely limited coded FOV. Radial hole masks [63], [64] have been considered to improve the coded FOV, however these approaches are still prone to artifacts for sources outside the FOV. Finally, the characteristic fluorescence photons ($58-85$ keV) from high-Z masks can obstruct low-energy reconstruction such as the 59 and 81 keV lines from ^{241}Am and ^{133}Ba, respectively [65].

Recent work has explored the replacement of passive mask materials with active (detector) elements to increase sensitivity, decrease weight, and increase the coded FOV. The detector material can attenuate gamma-rays but also actively reject Compton scattering interactions. Because the emphasis is placed on the masking using detector elements and not the opening between the elements, this concept is herein referred to as *active coded mask* imaging. This concept was first simulated with two coded planes of scintillator [17] and later demonstrated with the HEMI [18], [19], a compact hand-held two-plane CZT-based system (though using only one coded plane). Another advantage of the active coded mask design is that it facilitates Compton imaging of higher energy gamma-rays (hundreds of keV to a few MeV) that are of interest to nuclear security applications (e.g., ^{137}Cs at 662 keV, ^{238}U at 1001 keV). While this dissertation focuses only on the active coded mask imaging modality, a brief overview of Compton imaging principles and reconstruction techniques are given in the next section for completeness.

2.3.3 Compton Imaging

Coded aperture imaging is powerful when the photon energy is low because the modulation generates unique coding in the detected image. As the energy is increased, the photoelectric absorption cross-section decreases and the Compton scattering cross-section increases. The decrease in full-energy absorption and increase in scattering degrades the coding effect and results in blur. Gamma-rays can undergo Compton scattering in one detector, deposit a detectable amount of energy, and then eventually undergo a photoelectric absorption in another detector. While a variety of other scenarios exists (scatter-scatter-absorption, scatter-escape, pair produce-scatter-absorption, etc.), this section will focus only on two-interaction tracks. By measuring the interaction positions and energy depositions of the individual interactions, the photon path can be tracked through the system. Compton kinematics can then be used to determine a cone of incident angles from which the incident photon originated. The image is reconstructed by the overlap of multiple cones (in 3D) or circles (in 2D).

If the incident and scattered gamma-ray energies are known, the scattering angle can be solved with the Compton scattering formula

$$\mu = \cos\theta = 1 + \frac{m_e c^2}{E_\gamma} - \frac{m_e c^2}{E'_\gamma}\,. \tag{2.6}$$

The positions of the scattering interaction and the subsequent photoelectric absorption define the cone-axis and the scattering angle defines the cone opening angle. The cone of possible gamma-ray directions can then be back-projected into space from the scattering interaction location.

In practice, detectors have finite energy, position, and time resolution. First, finite time resolution will obscure the exact ordering of events. Coincidence gates must be

defined, which will introduce random coincidences into the data. Finite position resolution will introduce uncertainty into the cone axis. Typically, for a given interaction, the interaction positions are assigned to the center of the detectors. If the *lever arm*, or the distance between the two interactions, is small, this effect can become large. Finally, finite energy resolution will introduce uncertainty into the calculated scattering angle and limit the ability to determine if two events came from the same initial gamma-ray [66].

For a given two-interaction track, the measured energy depositions are used to attempt to sequence the events in time. First, the measured energy depositions in the detectors are assumed to sum to the incident gamma-ray energy. If $E_{tot} \leq 256$ keV, the kinematics of Eq. 2.6 states that the first interaction will always deposit less energy than the first. Therefore the sequence of the events can be known by comparing the energies. If $E_{tot} > 256$ keV, the sequencing will be ambiguous. Kinematic tests, such as the Compton Edge test can be used to discard forbidden sequences [66]. The Compton Edge refers to the maximum amount of energy that can be deposited for a given Compton scattering event. It occurs when the gamma-ray undergoes a complete backscatter ($\theta = \pi$). Each sequence is tested to determine whether the first energy deposition exceeds the Compton Edge. If so, the sequence is rejected.

In the far-field limit (parallel rays at infinity), each cone can be projected from the origin of the image space. The image space covers all of 4π and is defined as a unit sphere, $\mathbb{S}^2$. A unit vector is defined for the cone axis as $\hat{\vec{\omega}}$ and a point in $\mathbb{S}^2$ as $\hat{\vec{x}}$. The dot product of these two vectors then corresponds to the cosine of the angle between them. If this angle equals the scattering angle, then the point in $\mathbb{S}^2$ is on the cone. An infinitesimally thin cone back-projected into a discretized image space would result in pixellation effects. Therefore the cone is given a small width and intensity defined by a Gaussian function. The back-projected image is then the sum of all the cones

$$b(\hat{\vec{x}}) = \sum_{i=1}^{N} \frac{w_i}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(\hat{\vec{x}} \cdot \hat{\vec{\omega}} - \mu_i)^2}{2\sigma_i^2} \right), \tag{2.7}$$

where i runs over all the N cones, σ is the width of the cone, and w is the weight of the cone [67]. The width of the cone is selected to be smaller than the expected resolution of the system, as to be able to see resolution-based effects in the image reconstruction. One may be inclined to include the uncertainty in the scattering angle from the uncertainty in the measured energies into the cone width. For a small number of cones, this will certainly help with source location, however with a large number of cones, this will simply degrade the resolution because the uncertainty is introduced twice. Imagine a large number of cones with a small width being back-projected - the uncertainty in the cones position in space will be built into its projection and over a large number of cones, will define the resolution of the system. If the uncertainty is included, the

resulting image will be excessively blurred [67]. The weight of the cones are defined in two ways. In the case of ambiguous sequences, a normalized weight is applied to each sequence proportional to the Klein-Nishina differential scattering cross-section for the two scattering angles. The sequence with a larger probability of occurring is given a larger weight. The Klein-Nishina formula is given by

$$\frac{d\sigma}{d\Omega}(E_\gamma,\theta) = \frac{\alpha^2 r_c^2 P(E_\gamma,\theta)^2}{2}\left[P(E_\gamma,\theta) + P(E_\gamma,\theta)^{-1} - 1 + \cos^2(\theta)\right], \tag{2.8}$$

where α is the fine structure constant, r_c is the reduced Compton wavelength of the electron, and $P(E_\gamma,\theta)$ is given by

$$P(E_\gamma,\theta) = \frac{1}{1+\frac{E_\gamma}{m_e c^2}(1-\cos\theta)}. \tag{2.9}$$

For all cones, the lever arm L is also incorporated into the weight. Events with smaller lever arms are more likely to occur and will produce images with worse angular resolution. Therefore a weight of L^2 is applied in order to more heavily weight the larger lever arm events [67].

2.3.4 Far-field and Near-field

A distinction is made clear between imaging in the far- and near-field. Far-field imaging assumes the source standoff is large enough that the incident photon flux exhibits no divergence with respect to the size of the detector system (i.e., rays are parallel and the detector can be modeled as a point with no solid angle effects). In this case, the source can be represented as a distribution on $\mathbb{S}^2$, the unit 2-sphere, and thus the response measured in the detector depends only on the direction to the source. As the source is brought closer to the detector, effects from the finite size of the detector and beam divergence (i.e., magnification) must be accounted for. The near-field response is therefore dependent on the source position in $\mathbb{R}^3$.

In the coded mask imaging modality, the far-field assumption equates to an orthographic projection of the occluders onto the detectors. For example, with traditional passive coded apertures, this reduces the imaging problem to simply correlating the mask projection and the detected event pattern, where the mask projection is only a function of source direction. As near-field effects are introduced, the magnification of the mask pattern onto the detector plane adds an additional degree-of-freedom to the potential source locations. Therefore a perspective mask projection is required. However, if this response is known then 3D imaging is possible from a single static measurement.

In the Compton regime, the far-field approximation removes all solid angle effects and thus cones can be projected as if the detector was a point. In this case, the cones

are effectively projected as circles on $\mathbb{S}^2$. In the near-field, the solid angle effects can dominate and cones must be projected along the proper axes connecting the multi-site interactions. If the detector is large, it is possible for enough parallax to exist for projected cones to intersect in $\mathbb{R}^3$, facilitating static 3D imaging.

The distinction between near- and far-field is typically defined by the geometry of the detector and the angular resolution of the imaging modality [40]. The far-field assumption breaks down when the systematic image blur due to near-field effects is larger than the angular resolution of the system. The angular resolution of an occlusion based imager can be defined as

$$\delta\theta = \tan^{-1}\left(\frac{d}{L}\right), \tag{2.10}$$

where d is the opening distance (or aperture size) and L is the distance between the detector and occluder. The beam divergence from a near-field source is given by

$$\delta\theta = \tan^{-1}\left(\frac{r}{D}\right), \tag{2.11}$$

where r is the half-size of the detector and D is the source distance. By equating Eq. 2.10 and Eq. 2.11, one can solve for the source distance threshold between near- and far-field

$$D = \frac{r}{\tan(\delta\theta)}. \tag{2.12}$$

For example, the PRISM system described in Ch. 3 has a diameter of about 14 cm and a crystal separation of about 1.8 cm (center to center). The ideal angular resolution for this system is approximately 7.5° and the distinction between near- and far-field is made at a standoff of about 60 cm. Further discussion of near-field effects and the impact on 2D and 3D coded mask imaging will be presented in Sec. 4.3.

2.4 3D Imaging

Current state-of-the-art commercially available imagers [30], [31] and prototype imagers [8] are designed for static far-field data acquisition and image in 2D. While sometimes combined with a visual camera overlay for context, 2D images in many cases can result in ambiguity as to the true location of the source relative to the 3D scene (e.g., in or behind a box, in front or behind a wall). Furthermore, the static standoff measurements may require long dwell times (tens of min) to image weak or shielded sources. In other cases, 2D gamma-ray images are projected onto laser-based 3D models of the surrounding environment [68], though this approach still suffers from the same ambiguity problem when only one static gamma-ray measurement is performed. The introduction of unrestricted motion to these imaging systems can improve detection sensitivity while facilitating 3D gamma-ray imaging.

2.4.1 Contextual Sensing, SLAM and SDF

The use of SLAM for 3D radiation imaging was first demonstrated with red-green-blue-depth (RGB-D) Microsoft Kinect cameras and RGB-D SLAM [69] on a free-moving cart-based double-sided-strip HPGe detector [36] and a hand-held dual-plane CdZnTe detector array [19]. However, the active infrared depth sensors on the Kinect have a range of $4-6$ m, are limited to indoors mapping scenarios and have a small field-of-view ($47^\circ \times 53^\circ$). Other work has demonstrated the use of light detection and ranging (LiDAR) sensors for indoor, outdoor, and wide-area ($10-100$ m) SLAM on a variety of manned and unmanned platforms [70], [71]. All the previous literature has focused on the Compton imaging modality for gamma-ray energies from several hundred keV to a few MeV. One of the primary goals of this work is to extend this capability to lower energies, using the active coded mask imaging modality.

The work of [19], [36] demonstrated the 3D imaging capability in real-time with a technique referred to as scene data fusion (SDF). Real-time reconstruction, providing a 3D scene and gamma-ray model to the user during a measurement, is critical in radiological source search as it provides feedback to the user and facilitates course correction, decreasing the time to detect and image suspected sources. The SDF approach is utilized here.

In this work, the contextual sensor array includes a LiDAR (Velodyne Puck Lite [72]) and an inertial measurement unit (IMU) (VectorNav VN-200 [73]). The data streams are processed using the Google Cartographer package [74] to produce pose estimates and a 3D scene model at a rate of $\sim$10 Hz. The list-mode data are time-correlated with the pose estimates to provide the position and orientation of the detection system in the global image space at every event. The image space is defined with a uniformly voxelized grid based on the extent of the measurement path. The bounds of the grid are set at 5 m from the path extremes in each spatial dimension. This distance represents a balance between creating a large image space and maintaining appreciable sensitivity in every voxel. The choice of voxel size also presents a trade-off between the spatial resolution of the image space and the gamma-ray reconstruction time (which scales with the number of voxels).

The point cloud (i.e., the collection of 3D points from the SLAM-aligned laser-scans) is used to generate an occupancy grid over the voxel space, returning only the voxels that contain points and that have neighboring voxels. To reduce noise from spurious points in the cloud, each voxel is required to contain at least 10 points and have at least 4 neighbor voxels. The gamma-ray reconstruction is then constrained to the occupied voxels, limiting the source distribution to the surfaces of objects in the scene in which the LiDAR measured a reflection. The voxel constraint can improve image accuracy and decrease noise under the assumption that gamma-ray sources are not present in free-space, and the overall reduction in the number of voxels can be substantial ($> 90\%$), significantly improving reconstruction speed. A threshold is set on the number of occu-

pied voxels (e.g., 10^5) to regulate the reconstruction time. If the threshold is crossed, the voxel size is increased. The voxel size is typically initialized to 10 cm.

In the current implementation, the image reconstruction assumes a dynamic detector system with respect to a stationary environment. Moving objects are neither tracked nor removed from the point cloud. Therefore the deterministic occupancy model described above is susceptible to noise from transient objects in the scene such as moving people or vehicles. Moreover, depending on the cuts used, points from the operator carrying the instrument can remain in the point cloud, leaving a trail of points near the measurement track. While not a significant issue for the measurements shown in this work, techniques such as those in [75] could be used to remove these points.

Since the SDF constraint currently limits the reconstruction to the point cloud (i.e., surfaces), voxels inside of closed objects (e.g., box, shipping container) or behind walls are labelled as free and not included in the reconstruction. This dual-state occupancy model (occupied or free) could be improved with a tri-state model (occupied, free, or unknown) to capture voxels inside of closed volumes or unexplored spaces. The reconstruction could then be performed over both occupied and unknown voxels to facilitate a full volumetric reconstruction. Open-source tools utilizing ray-casting techniques could be used to determine tri-state occupancy [76].

It is important to note that the gamma-ray reconstruction is currently limited by errors in the point cloud and pose uncertainties. It is expected that SLAM algorithms will continue to improve with developments in autonomous vehicles, however detailed characterizations of SLAM uncertainties will be necessary to understand their impact on free-moving gamma-ray image reconstruction and SDF. This is currently outside the scope of this work.

2.4.2 Visualization

Visualizing a high dimensional 3D data product including scene, pose, and gamma-ray information with enough contrast and little clutter can be challenging. This work utilizes three different methods of visualization, depending on the demonstration. First, the gamma-ray image is superimposed as a 3D contour plot on the 3D point cloud. The point cloud is colorized by the LiDAR return intensities (i.e., a measure of reflectivity) for added contrast. Multiple viewpoints of the scene are provided to give a sense of depth. However, generating and transmitting 3D images can be computational expensive and prohibit a real-time visualization framework. Therefore, in some cases, it is appropriate to show top-down 2D projection images on a downsampled point cloud. Finally, following the conclusion of a measurement, high-resolution 3D images can be produced with offline processing, in $\mathcal{O}$(min). Currently the point cloud is colorized by interpolating the gamma-ray intensity map and using open-source graphics software to render the final product. The lower 10% of the gamma-ray intensities are replaced with

the LiDAR return intensities to provide higher contrast. In the future, photogrammetry and structure-from-motion [77]–[81] can be used to generate colorized point clouds based on the RGB camera stream during the measurement.

2.5 Image Reconstruction

Gamma-ray imaging attempts to reconstruct the distribution of source activity (emission intensities) in some space given a set of discrete measurements from a static or dynamic gamma-ray sensitive detector or detectors. The sampling of the incident gamma-ray flux distribution (originating from the distribution of source activity) is called the "forward problem". The forward problem projects data from the source activity distribution space to the measured detector response space. The expectation of this projection is fully described by the system matrix and, as they involve the independent counting of discrete events, the measurements are Poisson samples of this projection. The process of reconstructing the source distribution in the image space from measured samples in the detector space is called the "inverse problem". This section sets out to describe several, but by no means all, techniques to solve the inverse problem.

First, some notation is defined. The discretized image space space is of size J and this space can either be 2D where the term pixel is used to describe a 2D section or simply a direction on $\mathbb{S}^2$, or in 3D where the term voxel is used to a describe a 3D volume of $\mathbb{R}^3$. The distribution of source intensities (i.e., the image) in units of activity (Bq) is therefore a $J \times 1$ matrix in the positive real space, expressed as $\boldsymbol{\lambda} \in \mathbb{R}_+^J$. A single element of the image space is j and the intensity in that element is therefore λ_j. The detector space describes the collection of data (measurements in units of counts per unit integration time) of size I from one or many detector elements at a discrete position in a discrete amount of time, expressed as $\boldsymbol{x} \in \mathbb{R}_+^I$. A single element in detector space is i and a single measurement is therefore x_i. The mapping between image space and detector space is described fully by the system matrix $\boldsymbol{V} \in \mathbb{R}_+^{I \times J}$. This matrix, in units of inverse activity (Bq^{-1}), describes the geometric and detector efficiency of the I measurements relative to J image elements (i.e., V_{ij} is the probability of a photon emitted from image element j being detected in measurement i)

$$V_{ij} \propto \eta_{ij} t_i |\vec{r}_i - \vec{r}_j|^{-2}, \tag{2.13}$$

where η_{ij} is the angular response of measurement i to image element j, t_i is the integration time of measurement i, and $\vec{r}_i$ and $\vec{r}_j$ are the global position coordinates associated with measurement i and image element j, respectively. The angular response is determined either through measurement or simulation.

Next, it is useful to represent the gamma-ray imaging inverse problem using Bayes' Theorem

$$p(\boldsymbol{\lambda}|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\boldsymbol{\lambda})p(\boldsymbol{\lambda})}{p(\boldsymbol{x})}\,, \tag{2.14}$$

where $p(\boldsymbol{\lambda}|\boldsymbol{x})$ is the conditional probability of $\boldsymbol{\lambda}$ given $\boldsymbol{x}$ to be true, $p(\boldsymbol{x}|\boldsymbol{\lambda})$ is the conditional probability of $\boldsymbol{x}$ given $\boldsymbol{\lambda}$ to be true, and $p(\boldsymbol{\lambda})$ and $p(\boldsymbol{x})$ are marginal probabilities (i.e., the probability of observing $\boldsymbol{\lambda}$ and $\boldsymbol{x}$ independently of each other). Often times this expression is rewritten using logarithms, with each term following a specific naming convention

$$\log \underbrace{p(\boldsymbol{\lambda}|\boldsymbol{x})}_{\text{posterior}} = \log \underbrace{p(\boldsymbol{x}|\boldsymbol{\lambda})}_{\text{likelihood}} + \log \underbrace{p(\boldsymbol{\lambda})}_{\text{prior}} - \log \underbrace{p(\boldsymbol{x})}_{\text{evidence}}\,, \tag{2.15}$$

where the goal of the image reconstruction is to find the $\boldsymbol{\lambda}$ that maximizes the posterior distribution. The likelihood describes the probability of the measuring the detector data $\boldsymbol{x}$ from the image model $\boldsymbol{\lambda}$ (i.e., the forward problem), the prior encompasses any information known about the source distribution (e.g., sparsity or smoothness), and the evidence is independent of any model $\boldsymbol{\lambda}$ and is thus ignored in the optimization problem over $\boldsymbol{\lambda}$. Often times no prior information is assumed about source distribution, in which case, the maximization of the posterior is equivalent to maximizing the likelihood.

2.5.1 Poisson Likelihood

Gamma-ray measurements are governed by Poisson statistics since they involve independent counting of discrete events. The negative log-likelihood of a set of Poisson distributed measurements $\boldsymbol{x}$ from mean-rates $\bar{\boldsymbol{x}}$ is

$$\ell(\boldsymbol{x}|\bar{\boldsymbol{x}}) = [\bar{\boldsymbol{x}} - \boldsymbol{x} \odot \log \bar{\boldsymbol{x}} + \log[\Gamma(\boldsymbol{x}+1)]]^T \cdot \mathbf{1}\,, \tag{2.16}$$

where $\odot$ denotes element-wise multiplication and $\Gamma(\cdot)$ is the gamma function. In the absence of background, the mean-rates in units of counts per unit integration time are the forward-projection of the image

$$\bar{\boldsymbol{x}} = \mathbf{V} \cdot \boldsymbol{\lambda}\,, \tag{2.17}$$

and the measurements $\boldsymbol{x} \sim \text{Poisson}(\bar{\boldsymbol{x}})$. With a constant background b, the mean-rates take the form

$$\bar{\boldsymbol{x}} = \boldsymbol{V} \cdot \boldsymbol{\lambda} + b\boldsymbol{t}\,, \tag{2.18}$$

where $\boldsymbol{t} \in \mathbb{R}_+^I$ are measurement time durations. Detection algorithms that operate in variable background environments are currently under development and could eliminate the need for the constant background assumption [82].

The notation presented here is for a single detector system, however, the formulation is easily extensible to multi-detector systems. In this case, the number of measurements I will increase according to the number of detectors D (i.e., $I \leftarrow I \times D$). Additionally, the background rate in each detector is treated as a free variable, such that $b \Rightarrow \boldsymbol{b} \in \mathbb{R}_+^D$ and $\boldsymbol{t} \Rightarrow \boldsymbol{t} \in \mathbb{R}_+^{I \times D}$, where only a single element of each row in $\boldsymbol{t}$ is non-zero and Eq. 2.18 takes the form $\bar{\boldsymbol{x}} = \boldsymbol{V} \cdot \boldsymbol{\lambda} + \boldsymbol{t} \cdot \boldsymbol{b}$.

2.5.2 Back-projection

Equations such as Eq. 2.17 can be solved for $\boldsymbol{\lambda}$ with a direct inversion

$$\hat{\boldsymbol{\lambda}} = \boldsymbol{V}^{-1} \cdot \boldsymbol{x}\,. \tag{2.19}$$

Without noise ($\boldsymbol{x} \approx \bar{\boldsymbol{x}}$) and an invertible $\boldsymbol{V}$, this solution is exact. However, in the case presented here, $\boldsymbol{V}$ is typically sparse and non-square and therefore rarely invertible or well-conditioned enough for a reliable pseudo-inverse operation. Also in the presence of noise (in $\boldsymbol{x}$) and uncertainty (in $\boldsymbol{V}$), the multiplication of $\boldsymbol{V}^{-1}$ will produce significant noise amplification and skew. An alternative to direct inversion is simple back-projection (SBP), replacing $\boldsymbol{V}^{-1}$ with $\boldsymbol{V}^T$

$$\hat{\boldsymbol{\lambda}} = \boldsymbol{V}^T \cdot \boldsymbol{x}\,. \tag{2.20}$$

For every measured event, this approach adds intensity to every image element proportional to the probability of detecting the event from a photon emitted from that image element. To account for the measurement to be more sensitive to some image elements than others, a sensitivity correction step is typically used

$$\hat{\boldsymbol{\lambda}} = (\boldsymbol{V}^T \cdot \boldsymbol{x}) \oslash \boldsymbol{\varsigma}\,, \tag{2.21}$$

where $\oslash$ denotes element-wise division and the sensitivity is defined as

$$\boldsymbol{\varsigma} = \boldsymbol{V}^T \cdot \boldsymbol{1}\,. \tag{2.22}$$

Other corrections, or filters, exist with the ramp filter being a well-known filter in medical projection imaging [83]. The filtering approach is appropriately named filtered back-projection (FBP) in the literature. In the case presented here, if Eq. 2.17 is solved using least-squares

$$\underset{\boldsymbol{\lambda}}{\operatorname{argmin}}\ (\boldsymbol{x} - \boldsymbol{V} \cdot \boldsymbol{\lambda})^T (\boldsymbol{x} - \boldsymbol{V} \cdot \boldsymbol{\lambda})\,, \tag{2.23}$$

the resultant solution takes the form

$$\hat{\boldsymbol{\lambda}} = (\boldsymbol{V}^T \boldsymbol{V})^{-1} \cdot \boldsymbol{V}^T \cdot \boldsymbol{x}\,, \tag{2.24}$$

which can be viewed as a simple back-projection followed by a 2D ramp filter. However, for this specific filter, least-squares inherently assumes Gaussian statistics and may not

perform well in count limited scenarios (which can be the case in the search scenario). Furthermore, this filtering technique can produce negative values in the image, losing count preservation (i.e., $[\boldsymbol{V} \cdot \hat{\boldsymbol{\lambda}}] \cdot \mathbf{1} = \boldsymbol{x} \cdot \mathbf{1}$) and some degree of interpretability in the image.

While these analytical solutions are fast to compute, they tend to produce low-quality (i.e., blurry) images and can be difficult to interpret. Iterative techniques can be used to produce higher quality, count-preserved images, however, at the cost of added computational complexity.

2.5.3 Maximum Likelihood Expectation Maximization

Maximum likelihood expectation maximization ML-EM is an statistically founded and highly general iterative algorithm that solves for the source intensities $\boldsymbol{\lambda}$ and background b that maximum the Poisson likelihood, by minimizing Eq. 2.16

$$\hat{\boldsymbol{\lambda}}, \hat{b} = \underset{\boldsymbol{\lambda}, b}{\operatorname{argmin}}\ \ell(\boldsymbol{x}|\boldsymbol{\lambda}, b)\,. \tag{2.25}$$

In this case, $\hat{\boldsymbol{\lambda}}$ is referred to as the maximum likelihood estimate. Without background, the update equation at iteration $q+1$ is

$$\hat{\boldsymbol{\lambda}}^{q+1} = (\hat{\boldsymbol{\lambda}}^{q} \oslash \boldsymbol{\varsigma}) \odot \boldsymbol{V}^T \cdot [\boldsymbol{x} \oslash (\boldsymbol{V} \cdot \hat{\boldsymbol{\lambda}}^{q})]\,. \tag{2.26}$$

The convergence of ML-EM is not dependent on the initial estimate, thus the source intensities are typically initialized with a flat image ($\hat{\boldsymbol{\lambda}}^{q} = \mathbf{1}$). With the inclusion of a constant background rate across all measurements, the updates can be separated into source intensity and background equations. First defining the comparator term as

$$\boldsymbol{\xi}^{q} = \boldsymbol{x} \oslash (\boldsymbol{V} \cdot \hat{\boldsymbol{\lambda}}^{q} + \hat{b}^{q}\boldsymbol{t})\,, \tag{2.27}$$

the updates become

$$\begin{aligned} \hat{\boldsymbol{\lambda}}^{q+1} &= (\hat{\boldsymbol{\lambda}}^{q} \oslash \boldsymbol{\varsigma}) \odot (\boldsymbol{V}^T \cdot \boldsymbol{\xi}^{q})\,, \\ \hat{b}^{q+1} &= \frac{\hat{b}^{q}}{T}(\boldsymbol{t}^T \cdot \boldsymbol{\xi}^{q})\,, \end{aligned} \tag{2.28}$$

where the total measurement time $T = \boldsymbol{t}^T \cdot \mathbf{1}$ and $\hat{b}^0$ is typically the median of $\boldsymbol{x}$.

The number of iterations used in ML-EM represents a balance between contrast recovery and image noise amplification [84]. A statistical stopping criterion can be used (e.g., a predefined tolerance on the change in the Poisson likelihood), however the likelihood monotonically increases and thus does not translate directly to an image quality metric in terms of noise or artifacts [85]. In many cases, the number of iterations is set to an arbitrary number based on past experience or, in the some cases presented here, reconstruction speed requirements.

2.5.3.1 List-mode ML-EM

The computational complexity of each ML-EM iteration is $\mathcal{O}(I \times J)$. In the "bin-mode" formulation [38], the dimension I is of size $P \times D$ where P is the number of poses and D is the number of detectors. A typical free-moving measurement scenario can contain 10^3 to 10^4 poses, therefore I can easily be on the order of 10^5 to 10^6 when using high-dimensional detector arrays. In most cases, the number of measured gamma-ray counts N in a particular ROI is significantly smaller than $P \times D$. Therefore a "list-mode" formulation of ML-EM [39], in which the dimension I is of size N, is more appropriate for this type of imaging problem.

The list-mode data $\boldsymbol{y} \in \mathbb{R}_+^N$ consists of N events, where each event y_n represents a single measured interaction in detector $d \in \mathbb{Z}_+^D$ at pose $p \in \mathbb{Z}_+^P$ with energy deposition E in some ROI (e.g., 122 ± 10 keV). The $\{n, j\}$ element of the list-mode system matrix $\boldsymbol{K} \in \mathbb{R}_+^{N \times J}$ represents the probability of a gamma-ray emitted from image element j producing event y_n. Without background, the list-mode ML-EM update equation (at iteration $q+1$) is given by

$$\hat{\boldsymbol{\lambda}}^{q+1} = (\hat{\boldsymbol{\lambda}}^q \oslash \boldsymbol{\varsigma}) \odot \boldsymbol{K}^T \cdot [\mathbf{1} \oslash (\boldsymbol{K} \cdot \hat{\boldsymbol{\lambda}}^q)]\,, \tag{2.29}$$

where the sensitivity is still computed over the complete system matrix $\boldsymbol{V}$ (all possible events $P \times D$, not just the N events measured in $\boldsymbol{y}$), $\boldsymbol{\varsigma} = \boldsymbol{V}^T \cdot \mathbf{1}$.

2.5.4 Maximum A Posteriori

No prior information about the source intensities distribution was assumed in the traditional ML-EM approach detailed above. *A priori* knowledge about the image distribution can be incorporated into Eq. 2.25 with

$$\hat{\boldsymbol{\lambda}}, \hat{b} = \underset{\boldsymbol{\lambda}, b}{\operatorname{argmin}}\ \ell(\boldsymbol{x}|\boldsymbol{\lambda}, b) + \rho R(\boldsymbol{\lambda})\,, \tag{2.30}$$

where R is a convex regularizer function or penalty on the source intensities, ρ controls the strength of the regularization, and $\hat{\boldsymbol{\lambda}}$ is known as the maximum *a posteriori* (MAP) estimate.

One method that attempts to solve this problem is Green's one step late (OSL) MAP-EM algorithm [86] that uses the iterative updates

$$\begin{aligned} \hat{\boldsymbol{\lambda}}^{q+1} &= [\hat{\boldsymbol{\lambda}}^q \oslash (\boldsymbol{\varsigma} - \rho \boldsymbol{D}(R(\hat{\boldsymbol{\lambda}}^q))] \odot (\boldsymbol{V}^T \cdot \boldsymbol{\xi}^q)\,, \\ \hat{b}^{q+1} &= \frac{\hat{b}^q}{T}(\boldsymbol{t}^T \cdot \boldsymbol{\xi}^q)\,, \end{aligned} \tag{2.31}$$

where $\boldsymbol{D}$ is a vector of partial derivative of the elements of $\boldsymbol{\lambda}$. The OSL term is used here as the partial derivative is applied to the estimate of the previous step. Another,

simpler, method is to minimize the two components of Eq. 2.30 in two separate steps (i.e., perform an operation f on the image estimate $\hat{\boldsymbol{\lambda}}^q \xrightarrow{f} {}^*\hat{\boldsymbol{\lambda}}^q$ between every iteration of ML-EM).

2.5.4.1 Regularization

Radiation source distributions are generally expected to be smoothly varying in space, apart from sharp discontinuities in spatial extent due to either the nature of the source itself or the surrounding environment. Several de-noising techniques exist to remove unphysical high-frequency noise, however these also tend to smooth out or blur important edges in the image space. Total variations regularization is a common approach to remove high-frequency detail from images while preserving edges [87]. In the 2D formulation (where the image element index j is broken into two components k and l), the TV regularizer can be written as

$$R_{\mathrm{TV}}(\boldsymbol{\lambda}) = \sum_{k,l} \sqrt{(\lambda_{k+1,l} - \lambda_{k,l})^2 + (\lambda_{k,l+1} - \lambda_{k,l})^2 + \epsilon} \tag{2.32}$$

where the artificial parameter ϵ is required to ensure differentiability. The Python scikit-image library also contains a TV regularizer function that can be used to perform TV de-noising on an image of N-dimensions [88].

Furthermore, Candès et al. [89] have shown that, under certain conditions, a function can be recovered perfectly even with a sampling rate less than the Nyquist-Shannon sampling theorem. This is referred to a compressive, or compressed, sensing. The two major requirements for compressive sensing are incoherence in the sampling and the signal (the image in this case) must be sparse in some domain. Incoherence is usually satisfied with the combination of irregular pose sampling from free-moving systems and active coded masking of the detector elements. Sparsity can be achieved in the wavelet domain, where images tend to have sparse wavelet coefficients [90]. The wavelet transform decomposes images into frequency components while retaining spatial information, unlike the Fourier transform. Wavelet de-noising can then zero out the coefficients under some threshold to remove noise while conserving high-frequency features such as discontinuities (i.e., edges). The non-shift invariant nature of the wavelet transform is handled using the cycle spinning technique [91]. Wavelet de-noising and cycle spinning functions are also available in the Python scikit-image library for N-dimensional images.

Often times in radiological source search, the source distribution itself is known to be sparse (i.e., one or a few compact or point-like sources of radiation). Lingenfelter et al. [92] studied the effects of various penalties to enforce sparsity in the image intensity distributions. The sparsity-enforcing l_0 and l_1 norms were shown to produce solutions that were equivalent or scaled versions of the ML-EM solution. The authors proposed

a non-convex penalty based on the sum of the logarithm of all image intensity values resulting in a sparser solution than traditional ML-EM:

$$R_{\log}(\boldsymbol{\lambda}) = \log\left(\frac{\boldsymbol{\lambda}}{\delta} + 1\right) \cdot \mathbf{1}\,, \tag{2.33}$$

where δ is a scale parameter.

Another prior used in image reconstruction, developed primarily for emission and transmission tomography [93], [94], is the gamma prior

$$R_{\Gamma}(\boldsymbol{\lambda}) = (\boldsymbol{\lambda} \odot \boldsymbol{\alpha} \oslash \boldsymbol{\beta}) \cdot \mathbf{1} - (\boldsymbol{\alpha} - 1) \cdot \log \boldsymbol{\lambda}\,, \tag{2.34}$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}_{+}^{J}$ correspond to the mean (β_j) and variance (β_j^2/α_j) of the gamma density of image element j. This follows naturally for the underlying Poisson intensities as the gamma distribution is the conjugate prior of the Poisson likelihood. More importantly, a gamma distribution with a constant low mean and variance in each voxel will enforce sparsity as it penalizes the addition of source intensity far from the mean.

2.6 Reconstruction Acceleration Techniques

Real-time 3D imaging is critical to the radiological source search scenario in order to provide feedback to the operator and facilitate course correction to more efficiently localize sources. Iterative reconstruction techniques such as ML-EM can improve image quality, however, come at the cost of a higher computational and memory burden. Furthermore, the dimensionality of 3D image reconstruction with free-moving detector arrays can increase exceedingly quickly as gamma-ray data is collected and large spaces are explored. In this section, two methods for reconstruction acceleration for scalable real-time 3D imaging are described.

2.6.1 GPU Parallelization

The number of poses P and voxels J used in the reconstruction will increase as the measurement progresses and the system explores new space. The memory required to store the complete system matrix (of size $P \times D \times J$) can quickly exceed the available random access memory (RAM) on single-board computers onboard small free-moving systems. In this case, elements of the system matrix must be computed on the fly during reconstruction. This approach can take $\mathcal{O}(\mathrm{min})$ when performed on a low-parallelizable central processing unit (CPU). The use of a highly-parallelizable GPU can significantly increase computational efficiency, facilitating real-time, $\mathcal{O}(\mathrm{s})$, imaging.

The list-mode ML-EM algorithm described in Sec. 2.5.3.1 was parallelized in this work using the OpenCL framework [95] and the Python package pyopencl [96]. OpenCL

was chosen here as it can be run on CPUs, integrated graphics cards, and dedicated GPUs without any restrictions on hardware architecture (unlike e.g., CUDA [97], which is only compatible with NVIDIA products). While higher performing, dedicated GPUs tend to be large in size and have considerable power requirements, currently prohibiting their use on small free-moving systems, such as unmanned aerial systems (UAS) with limited payloads. Therefore, results shown in Ch. 4 utilize the stock integrated graphics card (Intel Iris Plus Graphics 650) on a single-board computer.

2.6.2 Ordered Subset Expectation Maximization

Hudson and Larkin [98] proposed an ordered subsets expectation maximization (OS-EM) algorithm that processes radiation data in subsets (blocks) within each iteration and demonstrated accelerated convergence over traditional ML-EM by a factor proportional to the number of subsets using Single Photon Emission Computed Tomography (SPECT) data. The method takes the "divide and conquer" approach and uses the reconstructed image from one subset of data as the input image to the next subset reconstruction. The subsets of data can be random, time-sequential, non-overlapping or cumulative, though, the choice of subset in one application may not work well in others. Typically the nature of that data provides insight into natural choices of subsets (e.g., single projections in SPECT can form successive subsets). However, the authors note that it is advantageous to select subsets in a balanced way so that pixel activity contributes equally to any subset.

The bin-mode formulation of the reconstruction (without background) follows the update equation

$$\boldsymbol{\lambda}^{q,g+1} = \left(\boldsymbol{\lambda}^{q,g} \oslash \left[\boldsymbol{V}^{g\,T} \cdot \mathbf{1}\right]\right) \odot \boldsymbol{V}^{g\,T} \cdot \left(\boldsymbol{x}^{g} \oslash \left[\boldsymbol{V}^{g} \cdot \boldsymbol{\lambda}^{q,g}\right]\right), \tag{2.35}$$

where g is the subset number and $\boldsymbol{V}^{g}$ is the system matrix for subset g. Notice the sensitivity correction is only over the current subset. This update equation iterates over all subsets, using the resultant image from one subset-iteration as the input image to the next. An OS-EM iteration q is defined as an update through all subset-iterations. Notice that if one subset exists with all the data, the update equation reduces to ML-EM.

As with traditional ML-EM, depending on the dimensions of the data (i.e., number of counts vs. number of poses and detectors) a list-mode formulation can result in faster reconstructions. The list-mode OS-EM formulation is described in [99] and takes the form

$$\boldsymbol{\lambda}^{q,g+1} = \left(\boldsymbol{\lambda}^{q,g} \oslash \left[\boldsymbol{V}^{T} \cdot \mathbf{1}\right]\right) \odot \boldsymbol{K}^{g\,T} \cdot \left(\mathbf{1} \oslash \left[\boldsymbol{K}^{g} \cdot \boldsymbol{\lambda}^{q,g}\right]\right), \tag{2.36}$$

where $\boldsymbol{K}^{g}$ is the list-mode system matrix for subset g. Notice, similarly to list-mode ML-EM, the sensitivity correction is over the complete system matrix. The list-mode formalism has also been implemented using GPUs [100].

The primary disadvantage of OS-EM is that it can suffer from convergence issues in the presence of noise (i.e., convergence is never achieved and the solution cycles through limit points) and is sensitive to the choice of subsets. Modifications have been developed for globally convergent OS-EM in bin-mode [101] by relaxing the number of subsets over time (i.e., slowly approaching the ML-EM approach) and list-mode [102] using regularization.

Chapter 3

PRISM

3.1 Introduction

In recent years, emphasis has been placed on the development and deployment of detection and imaging systems on a variety of free-moving platforms such as hand-held instruments and (un)manned ground-based and aerial-based vehicles for gamma-ray source localization and radiation distribution mapping. In the radiological source search scenario, where the goal is often to efficiently localize compact radioactive sources in cluttered environments, hand-held free-moving systems can overcome the inverse square law with the capability to move much closer to suspected source locations to find weak sources and can potentially circumvent complex shielding configurations by sampling many perspectives of the source [20]–[23].

As an improvement to the hand-held HEMI system, this work proposes a novel design by rearranging the detectors into an active coded spherical configuration, resulting in an omnidirectional FOV and isotropic efficiency for both coded aperture and Compton imaging. The Portable Radiation Imaging Spectroscopy and Mapping (PRISM) system is a hand-held, omnidirectional, dual-mode (coded and Compton), spherical active coded array ($\varnothing$14 cm) of 1 cm^3 CZT detectors. The system is equipped with an auxiliary sensor array (optical camera, LiDAR and IMU) for real-time contextual sensing and SDF. While the system is capable of Compton imaging, this work focuses solely on the active coded mask imaging modality.

This chapter describes the design, optimization and development of the PRISM system and is organized as follows. The hardware design of the first prototype system, herein referred to as PRISM-v0, is given in Sec. 3.2. The design upgrade to the second prototype system, herein referred to as PRISM-v1, is outlined in Sec. 3.3. Sec. 3.4 includes some comments on the use of PRISM-v0 and PRISM-v1 in the following sections and chapters. Far-field directional system response simulations are discussed in Sec. 3.5. Finally, the methodology to optimize the mask (i.e., the number and configuration of

elements) of 1 cm^3 CZT crystals in the sphere for omnidirectional active coded mask imaging is described in Sec. 3.6.

3.2 Prototype Design (PRISM-v0)

PRISM-v0 was designed as the successor to HEMI, leveraging the work on 1 cm^3 CZT crystals with CPG electrodes [53], [54] pioneered at LBNL. Similar to HEMI, each CPG-CZT detector crystal is connected to an application specific integrated circuit (ASIC) based on [103] to read out, process (i.e., collecting and non-collecting anode grid differencing with relative gain), shape and amplify the anode grid signal waveforms. PRISM-v0 was not designed to read out the cathode signals, though this was a design upgrade in PRISM-v1. The modules (crystal and ASIC) are packaged into modular snap-in Lexan casings (see Fig. 3.1).

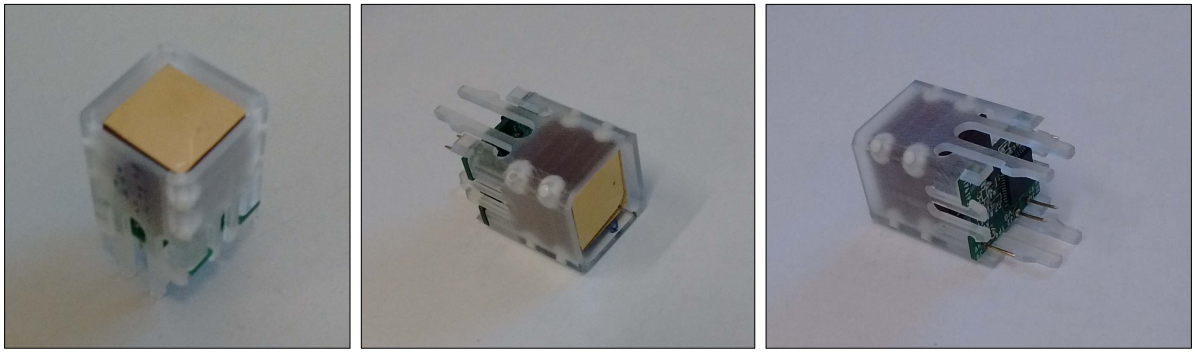

Figure 3.1. PRISM detector module comprising of 1 cm^3 CPG-CZT, ASIC and Lexan housing.

The detector array enclosure consists of six identical curved faces each with 32 detector slots (total of 192 available available detector locations) that come together to form a sphere. The detector modules are snapped into the interior of the enclosure and oriented such that the cathodes face in towards the center of the sphere. A flexible circuit runs along the exterior of the spherical enclosure and connects the individual ASIC boards to 12 analog-to-digital converter (ADC) cards (192 channels) on the signal acquisition motherboard (see Fig. 3.2). A total of 192 channels are available for processing the anode signals of all 192 detector slots, allowing for the number and configuration of the detector modules to change without modification to the electronics pipeline. The digitized signals are processed with a field programmable gate array (FPGA) on a digital signal processing board based on the Compact And Programmable daTa Acquisition Node (CAPTAN) architecture [104]. The processed signals are then routed via ethernet to a laptop computer carried in an accompanying backpack for further data processing, analysis and image reconstruction.

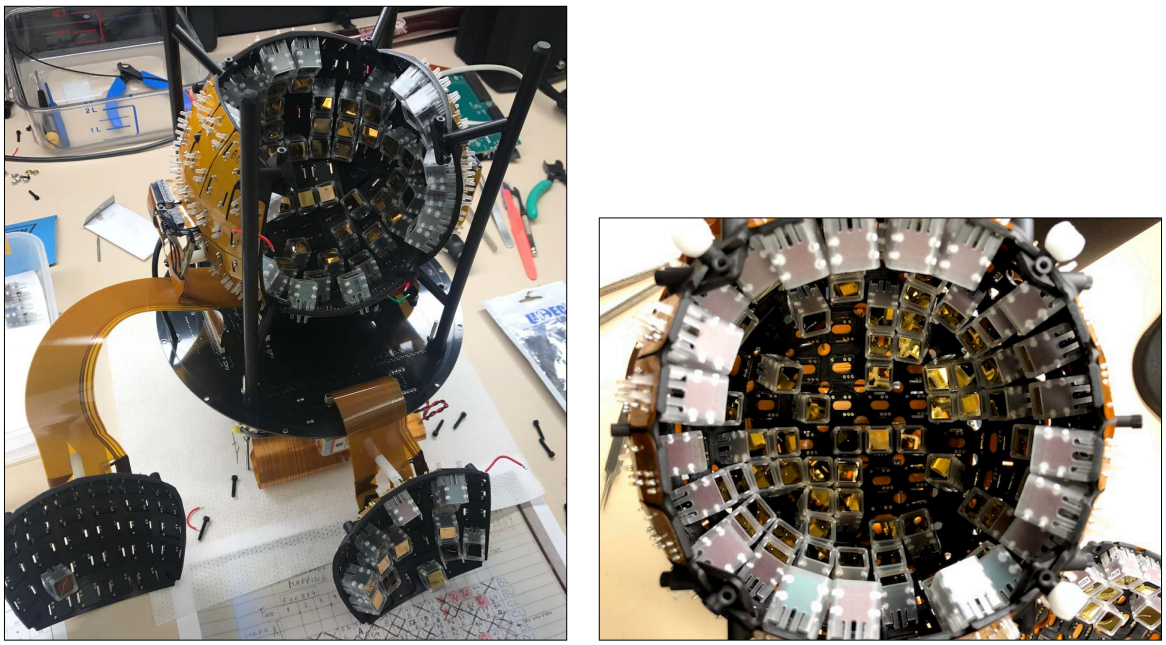

Figure 3.2. (Left) Assembly of PRISM-v0 spherical enclosure showing flex cable routing. (Right) Interior of populated sphere showing gold cathodes.

The auxiliary sensor array includes a LiDAR (Velodyne Puck Lite [72]) and IMU (VectorNav VN-200 [73]) for real-time contextual mapping and tracking. An optical camera was not included in the PRISM-v0 design, though it has been implemented in the PRISM-v1 design. Contextual data processing and SLAM (via the Google Cartographer package [74]) are also performed on the computer. A hand-held tablet (Microsoft Surface) is used for real-time remote control and visualization, connected to the laptop via USB. Two 98 W-hr Li batteries are used to power the detectors and contextual sensors. The computer is powered with its internal battery. The total operational battery life is $\sim$1 hr, due primarily to the demand placed on the computer (e.g., from image reconstruction and SLAM). With no heavy processing, the detectors can remain powered for $>$ 6 hr. The system can be powered on and begin acquiring data in $<$ 60 s. The hand-held device (excluding the background and tablet) weighs approximately 3.65 kg (with 96 detectors). Figure 3.3 shows the internal model and completely assembled PRISM-v0 prototype and Fig. 3.4 shows the entire user set-up including backpack and tablet.

Figure 3.5 shows some example singles gamma-ray spectra (individual channels and sum) for 30 min measurements of 10 μCi ^{133}Ba and ^{137}Cs. The sources were placed on-axis with 1 m standoff and the system was placed roughly 1 m above the floor. Figure 3.6 shows a picture of the source setup. Only 66 detectors were loaded in PRISM-v0 during this measurement. Energy calibration was done with a three-point linear fit using the ^{133}Ba (81 and 356 keV) and ^{137}Cs (662 keV) lines. The energy resolution of the detectors

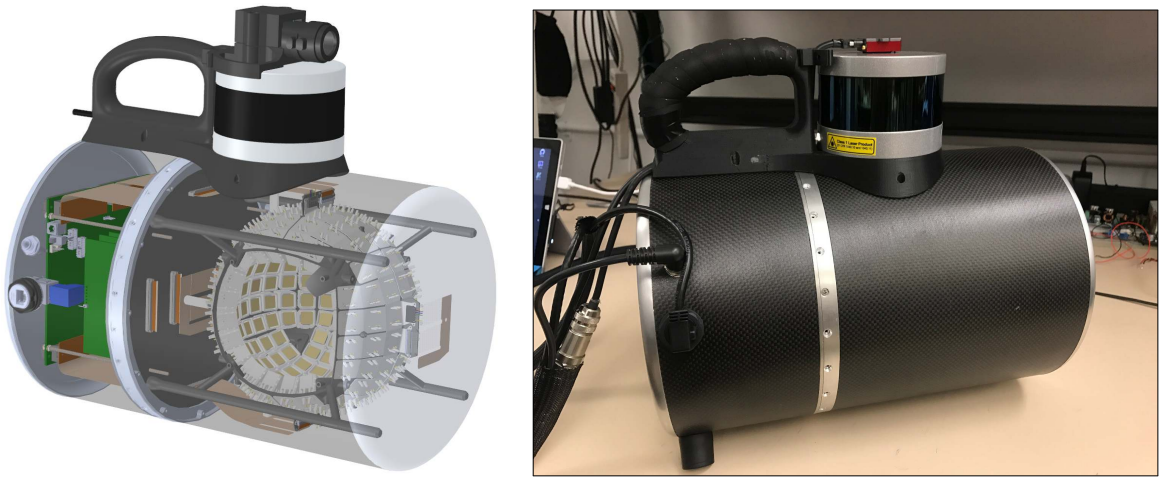

Figure 3.3. (Left) Model rendering of PRISM-v0 prototype showing the six modular faces each with 32 available detector slots. One face is removed to reveal the gold cathodes of the CPG-CZT detectors. (Right) Fully assembled prototype.

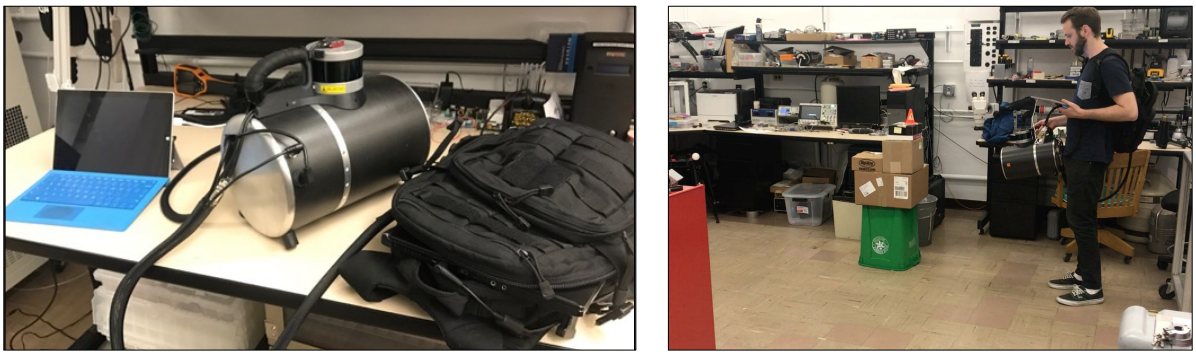

Figure 3.4. (Left) Setup of PRISM-v0 system including the hand-held detector, backpack holding batteries and computer, and hand-held tablet for control and visualization. (Right) Example user operating the system.

varied between 2−4% at 662 keV. Notice that some detectors display an attenuated 81 keV signature in the ^{133}Ba spectrum, due to occlusion. This effect is not significant at 662 keV due to the higher penetrating power of higher energy gamma-rays. This attenuation is what is used for active coded mask image reconstruction and thus demonstrates why the modality degrades at higher energies.

3.3 Design Upgrade (PRISM-v1)

The next iteration of the prototype system, PRISM-v1, improves upon the previous prototype in several ways. The contextual data processing and gamma-ray reconstruction

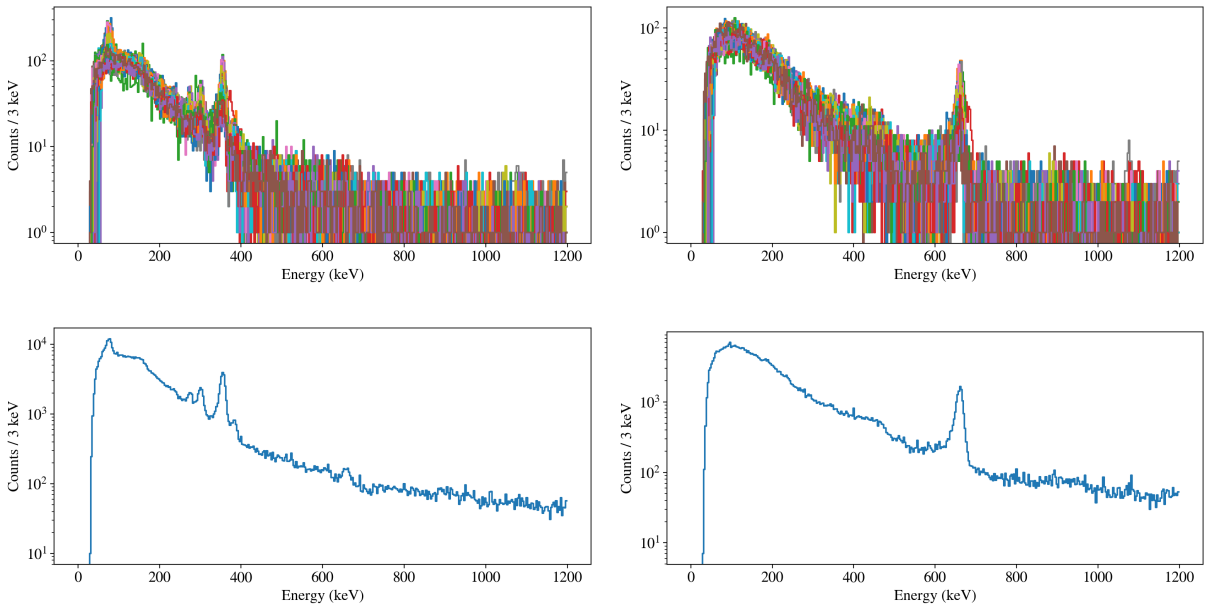

Figure 3.5. (Top) Individual channel energy-calibrated 30 min spectra of 10 μCi ^{133}Ba (left) and ^{137}Cs (right) placed at 1 m. A total of 66 detectors were loaded in the system at the time of the measurement. (Bottom) Total spectrum summed over all channels in top row.

Figure 3.6. Source and system setup for spectral measurement shown in Fig. 3.5

was moved to a single-board computer (Intel NUC7i7BNH [105]) and integrated directly into the detector enclosure, removing all external cables and the need for the backpack. A lightweight wireless tablet (iPad mini) is used for real-time remote control and visualization, connected to the system via WiFi. In addition to the same contextual sensors on PRISM-v0 (LiDAR and IMU), an optical camera (FLIR flea3 camera [106] with Kowa 3.5 mm/F1.4 lens [107]) is included for real-time augmented visualization. Several hardware components were upgraded to be smaller, lower-power, and lighter. A single 98 W-hr Li battery powers the entire system for $\sim$1 hr, though, again, the battery

life may fluctuate depending on the demand placed on the computer (e.g., from image reconstruction and SLAM). The system can be powered on and begin acquiring data in < 30 s. With the battery attached, the system weight is approximately 6.35 kg.

As mentioned in Sec. 2.2, the CPG electrode scheme is capable of determining event-by-event DOI in each CZT detector. PRISM-v1 makes use of this capability and reads out the cathode signal waveforms in addition to the anode grid. A flexible circuit still connects the individual ASIC boards to 12 ADC cards on the signal acquisition motherboard, now with 384 channels (both anode and cathode of all 192 detector locations). Figure 3.7 shows a diagram of all the upgraded components on PRISM-v1, Fig. 3.8 shows the components being assembled, and Fig. 3.9 shows the finished system and example use in the field.

Figure 3.10 shows example anode singles spectra (individual channels and sum) for a $\sim$10 min measurement of 20 μCi ^{133}Ba and ^{137}Cs. The sources were placed close to the detector enclosure and moved around all sides of the sphere in an attempt to uniformly illuminate all the detectors. Only 91 detectors were used in this measurement. Similar to PRISM-v0, the energy resolution of the detectors varied between 2$-$4% at 662 keV. However, some low-energy tailing is observed in the PRISM-v1 spectra. This is likely due to insufficient bias voltage applied to the detectors, leading to incomplete charge collection. Notice the 81 keV signature is not significantly different between channels in this case. This is because the sources were moved around the sphere, so occlusions have been averaged out.

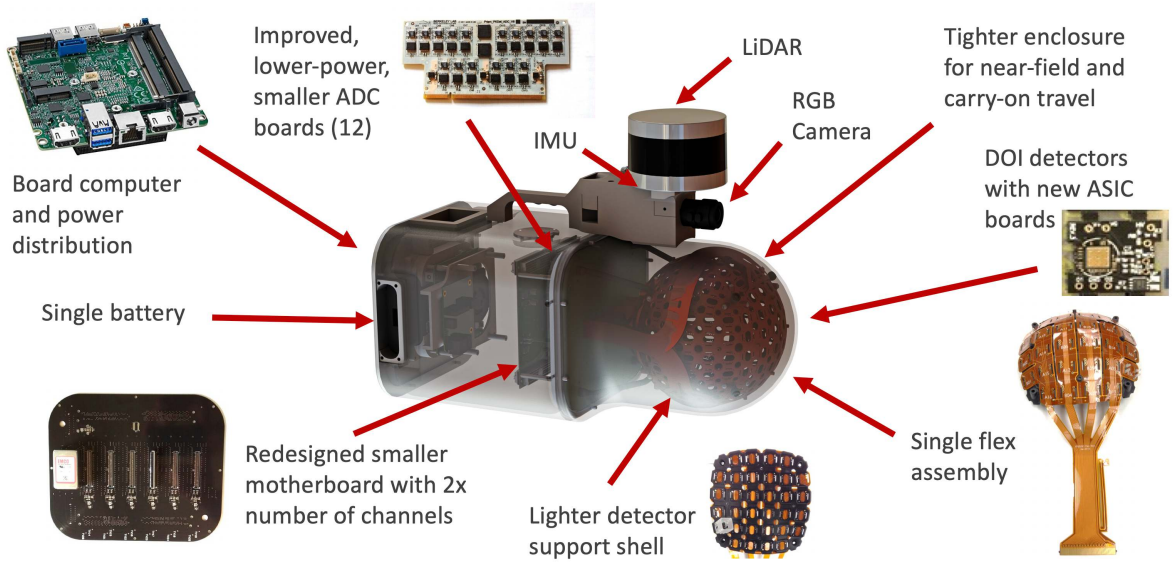

Figure 3.7. Diagram of upgraded components on PRISM-v1.

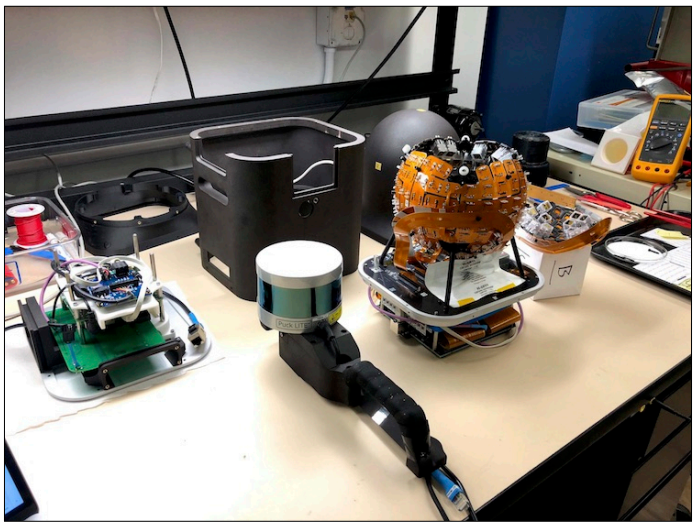

Figure 3.8. Assembly of PRISM-v1 components.

Figure 3.9. (Left) Fully assembled PRISM-v1 prototype system with accompanying wireless tablet for remote control and visualization. (Right) PRISM-v1 system being used in the field (literally).

3.3.1 Depth-of-interaction

PRISM-v1 facilitates the event-by-event read-out of deposited gamma-ray energy from both the anode grid and cathode electrodes. While the unipolar charge (i.e., electrons) sensing CPG anode electrode scheme produces accurate energy estimates regardless of depth (along the anode-cathode axis), the poor hole mobility in CZT causes the cathode events to suffer from depth-dependent incomplete charge collection. The cathode signal, however, can be used in an anode-cathode ratio to determine the DOI.

To demonstrate this capability, a single PRISM detector module was illuminated with a low-energy ^{241}Am source (60 keV) on the cathode, anode, and side of the crystal. A

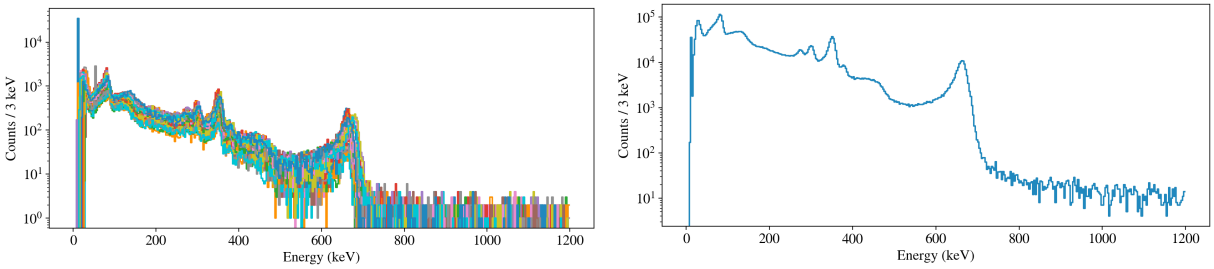

Figure 3.10. (Left) Individual channel and (right) total energy-calibrated spectra of 20 μCi^{133}Ba and ^{137}Cs moved continuously around the sphere to uniformly illuminate the detectors. A total of 91 detectors were loaded in the system at the time of the measurement.

finely collimated ^{241}Am source was also directed towards the middle of the side surface of the crystal. The mean free path of 60 keV gamma-rays is $<$ 1 mm in CZT and thus events can reasonably be assumed to occur on the surface of the crystal. Figure 3.11 shows the distribution of cathode-anode ratios for the four scenarios. A clear distinction is observed between the anode and cathode events and the uncollimated side-illuminated events produce a broad distribution of ratios, indicating events occurred along all depths. The collimated response indicates a $\sim$2 mm FWHM depth resolution for 60 keV events in the middle of the crystal. A uniform 2.5 mm resolution (i.e., four depth bins) is assumed in the rest of this work.

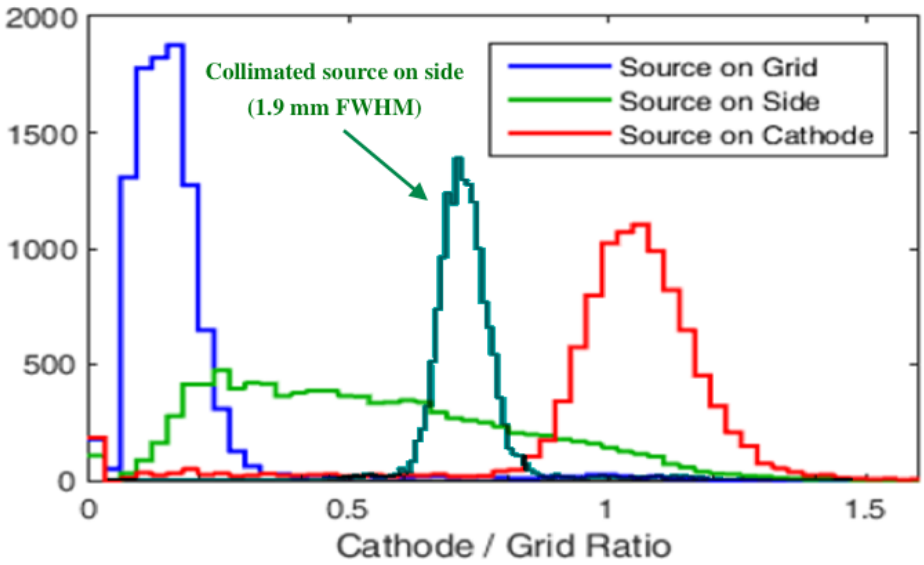

Figure 3.11. Cathode-to-anode-grid ratio distribution for an uncollimated ^{241}Am source on the cathode, anode and side of the crystal as well as a collimated ^{241}Am source source directed towards the middle of the side surface of the crystal (courtesy of P. Barton).

At the time of this writing, the coded configuration in PRISM-v1 consisted of 100 detectors. DOI capable module fabrication was still ongoing, and only 54 of the detectors were depth sensitive. Figure 3.12 shows the channel mapping unfolded in 2D, with the front and back pointing away and towards the electronics in the rear of the enclosure, respectively. Note the cathodes face in towards the center of the sphere and the anode grids face out. In this configuration, most of the depth-sensitive detectors are oriented such that the cathode faces are exposed to sources in front of the system.

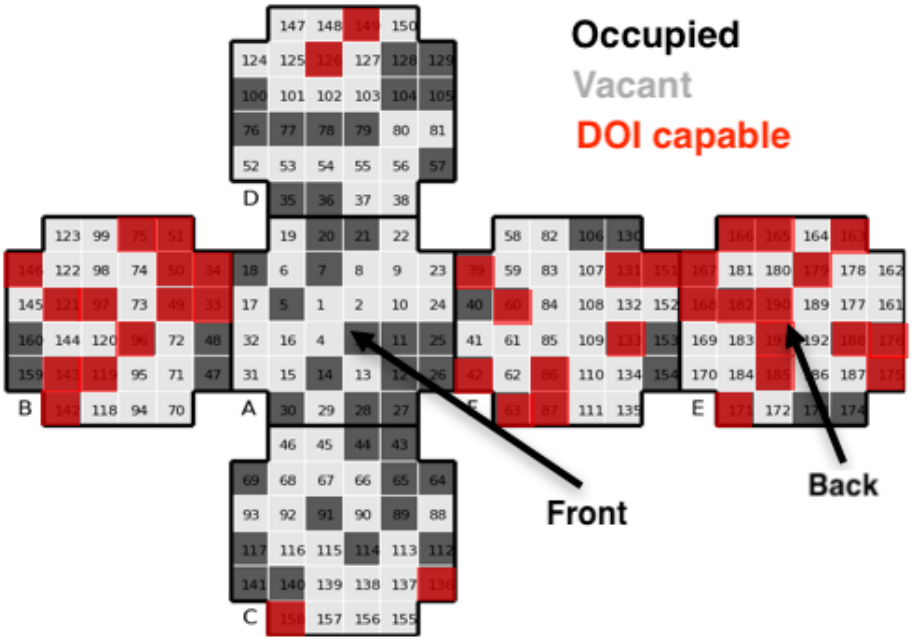

Figure 3.12. PRISM-v1 detector configuration unfolded in 2D, with populated channel marked in grey and DOI capable modules highlighted in red. In this diagram, the front and back point away and towards the electronics in the rear of the enclosure, respectively.

A simple measurement was performed to assess DOI performance in PRISM-v1. A 20 μCi ^{133}Ba source was placed 1 m in front of the sphere center, on axis. The measured singles energy spectrum is shown in Fig. 3.13 separately for the anode and cathode and in Fig. 3.14 as an anode-cathode 2D-histogram. The data show the inability of the cathode electrode to capture the full energy deposition and that the effect worsens with increasing energy.

Figure 3.15 shows the anode singles energy spectrum separated into the four separate depth bins. A clear separation between the anode and cathode bins at 81 keV is observed in the spectrum, indicating that most of the events do occur near the cathode, as expected for this configuration. The effect is less noticeable at higher energies due to the longer mean free path of gamma-rays distributing events more evenly throughout the crystal volumes as well as the degredation in DOI with increasing energy.

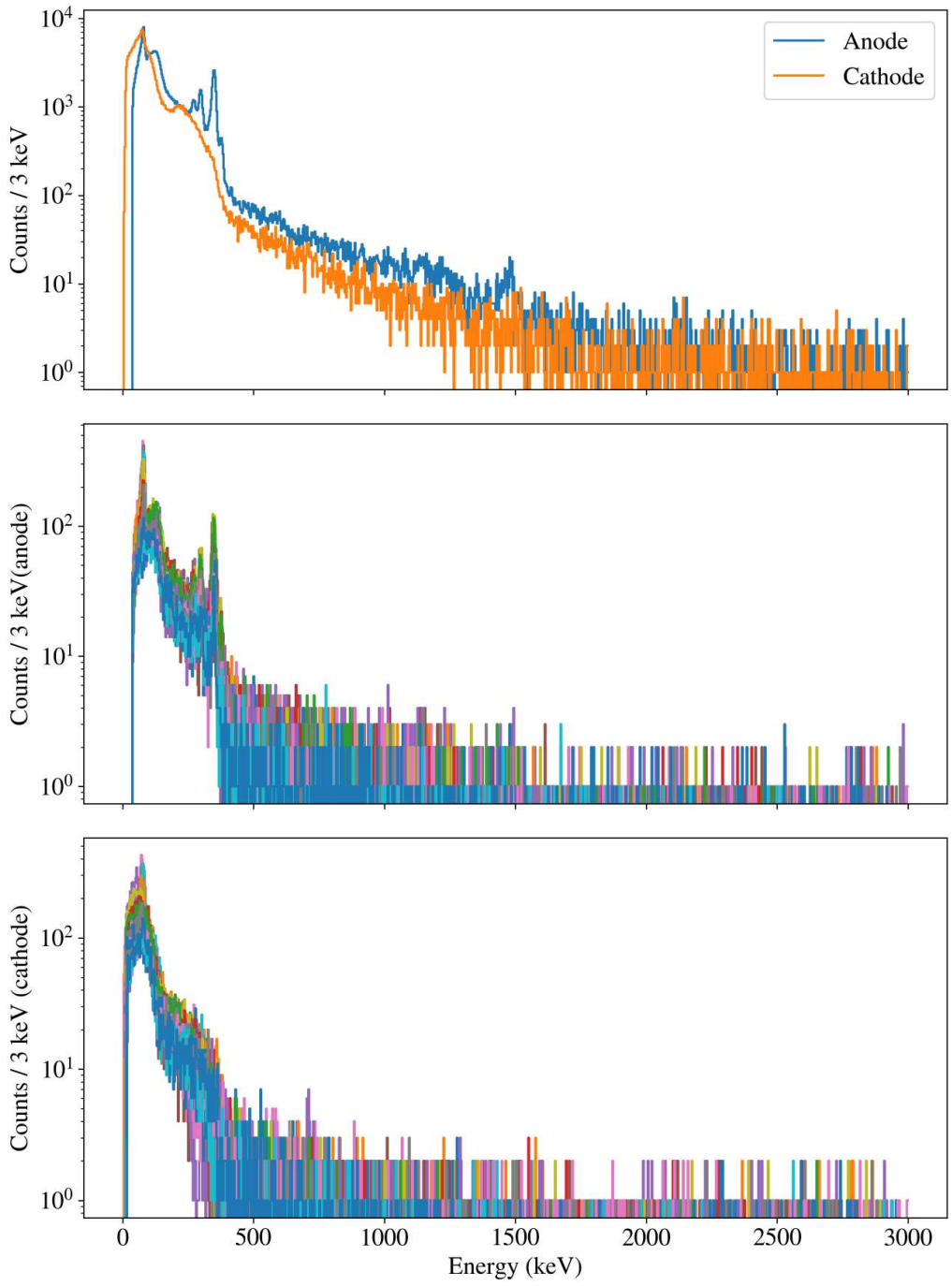

Figure 3.13. Anode and cathode energy spectra of a ^{133}Ba source placed 1 m in front of the sphere center, on axis.

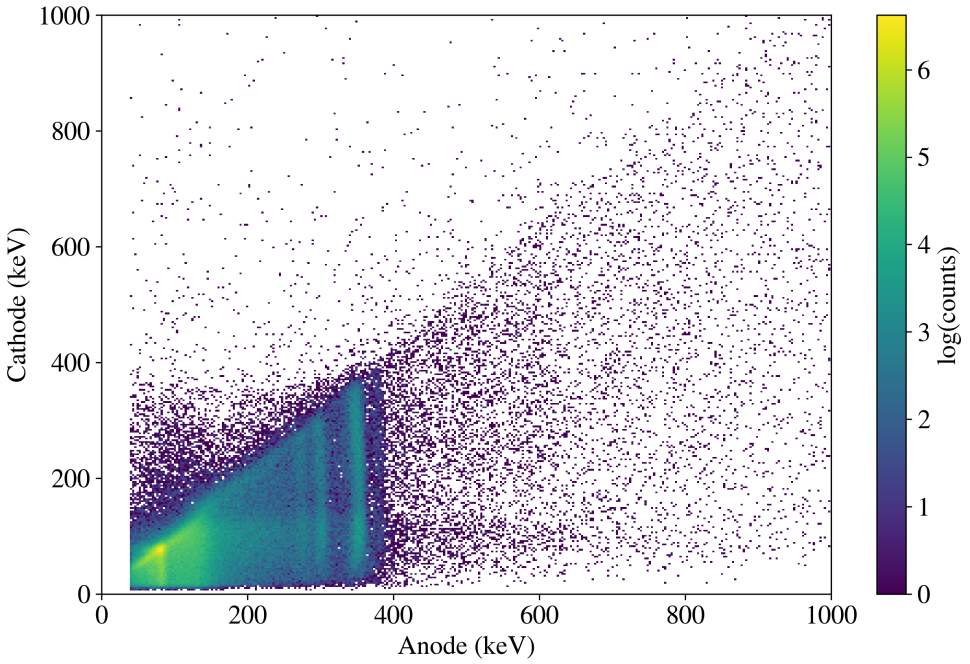

Figure 3.14. Anode-cathode 2D-histogram of the data in Fig. 3.13.

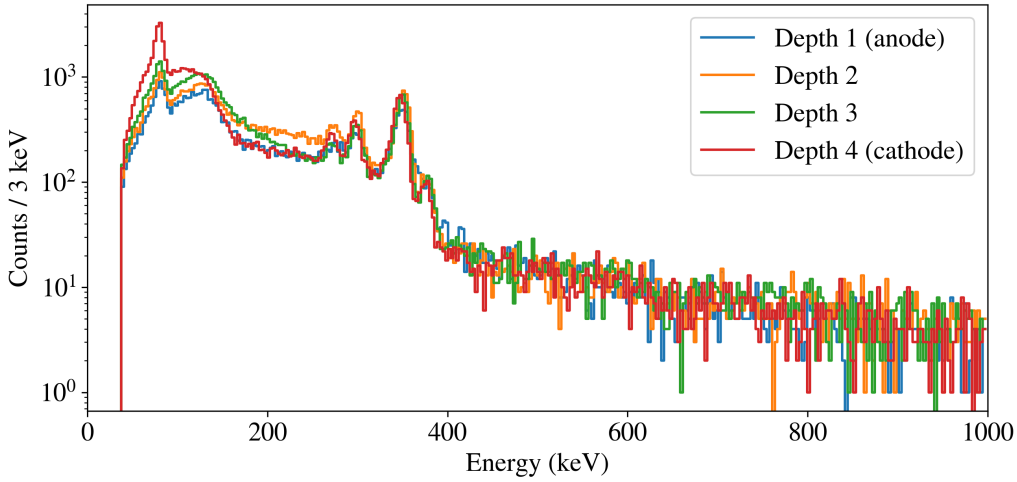

Figure 3.15. Anode energy spectrum of the data in Fig. 3.13 separated into the four depth bins.

3.4 Comments on PRISM-v0 and PRISM-v1

Two PRISM systems have been developed and used during the course of this work (see Fig. 3.16). Results will be shown from both systems in this dissertation, mentioning the specific system used where necessary. There are no fundamental differences in the

imaging capabilities between PRISM-v0 and PRISM-v1, as both utilized ∼100 detectors on a sphere. The primary distinction in imaging is the DOI capability in PRISM-v1, which will be treated separately.

Figure 3.16. PRISM-v0 and PRISM-v1 shown side-by-side.

3.5 Directional System Response

An element of the far-field directional system response, η_{ij}, is defined as the probability that a photon emitted from image pixel (direction) j is detected in detector i. The matrix, $\boldsymbol{\eta}$, or an estimate of it, is crucial to the active code mask image reconstruction process. In this work, two approaches were taken to determine $\boldsymbol{\eta}$.

In the first approach, a Monte Carlo simulation in Geant4 [108] was developed using a simplified model of the PRISM system represented only by the CZT crystals. In the low-energy regime of the active coded mask modality, surrounding materials including the contextual sensors, battery, electronics, and human operator will act as significant attenuators and down-scatterers and will add features to the directional response. Any systematic or statistical uncertainties in the response will propagate through reconstruction and manifest as artifacts or noise in the images. However, it is unclear as to where the largest uncertainties originate. Effects such as gain drift, detector noise, and pose uncertainty could certainly dominate the response uncertainty in real measurement scenarios, and the degree to which the geometry is accurately modeled in simulation could be second order. The results shown in Ch. 4 support this hypothesis, as high-quality images and accurate localizations were achieved with a simple simulation geometry.

The far-field image space pixellation was defined on the 2-sphere with the Hierarchical Equal Area isoLatitude Pixelation of a sphere (HEALPix) library [109] with $N_{\text{side}} = 16$ for a total of 3072 source directions (angular resolution of about 3.5°). When parallelized

on 768 CPUs using a high-performance computing cluster, the far-field system response of 10^6 system-incident gamma-rays from each of the 3072 source directions can be generated in about 90 s. The full-energy directional response for a random 100-detector configuration is shown in Fig. 3.17 for various energies, with and without DOI. A large response (yellow) indicates a high probability of full-energy absorption and a low response (purple) indicates a low probability of absorption (i.e., the detector is occluded). Both the detector and HEALPix indexing start at the north pole and spiral around the sphere towards the south pole. The DOI detectors follows the same indexing as the non-DOI detectors, but grouped into four distinct bins (i.e., detectors $0 \rightarrow 99$ and $300 \rightarrow 399$ correspond to the cathode side and anode side bins of all 100 detectors, respectively.)

In the second approach, a graphics-based program was developed using OpenGL to quickly generate the zero-energy (i.e., infinite attenuation and no scattering) directional system response. OpenGL is widely utilized for fast 2D and 3D graphics rendering using highly-parallelizable GPUs. Given a 3D model (i.e., cubic CZT crystals arranged in a sphere), OpenGL can quickly render an orthographic projection of the geometry (i.e., modeling a far-field point source with uniformly spaced mono-directional system-incident rays) including occlusion. In this approach, each detector crystal is indexed by a unique 8-bit red color value and DOI is included using a gradient of 8-bit blue values along the anode-cathode axis of each detector. Figure 3.18 shows a snapshot of the OpenGL display of the PRISM sphere with and without the DOI color schemes. For a given projection (i.e., source direction), the red and blue channels in each pixel in the display correspond to a detector index ($1 \rightarrow 192$) and a DOI value ($1 \rightarrow 256$), respectively. The blue values can be binned according to the depth resolution of the detector. Pixels are then histogrammed by color to determine the visible surface area (i.e., effective zero-energy flux) of each detector bin. The system response from all 3072 source directions using a 10^6 pixel display can be generated in under 25 s on a single NVIDIA Quadro K1100M graphics card.

This approach is useful in the case where a quick estimate of the directional system response is required (e.g., the mask optimization described in Sec. 3.6). While a zero-energy system response is unphysical, the results should not differ significantly from those for low-energy photons. At 60 keV, the mean free path in CZT is approximately 0.26 mm (2.6% of the detector thickness) and the ratio of the photoelectric absorption and Compton scattering cross-sections is > 50. Figure 3.19 shows the directional responses with and without DOI of the same random 100-detector configuration as above simulated with OpenGL. As in Fig. 3.17, the high and low end of the color scale represent large and small absorption probabilities, respectively. Notice how these responses do not appear significantly different than the 60 keV response simulated with Geant4.

Note that once an optimal mask configuration has been found, the Geant4 simulation described above could be run at several source energies to estimate the directional and energy dependence of the system response. A ray-tracing approach could also be used

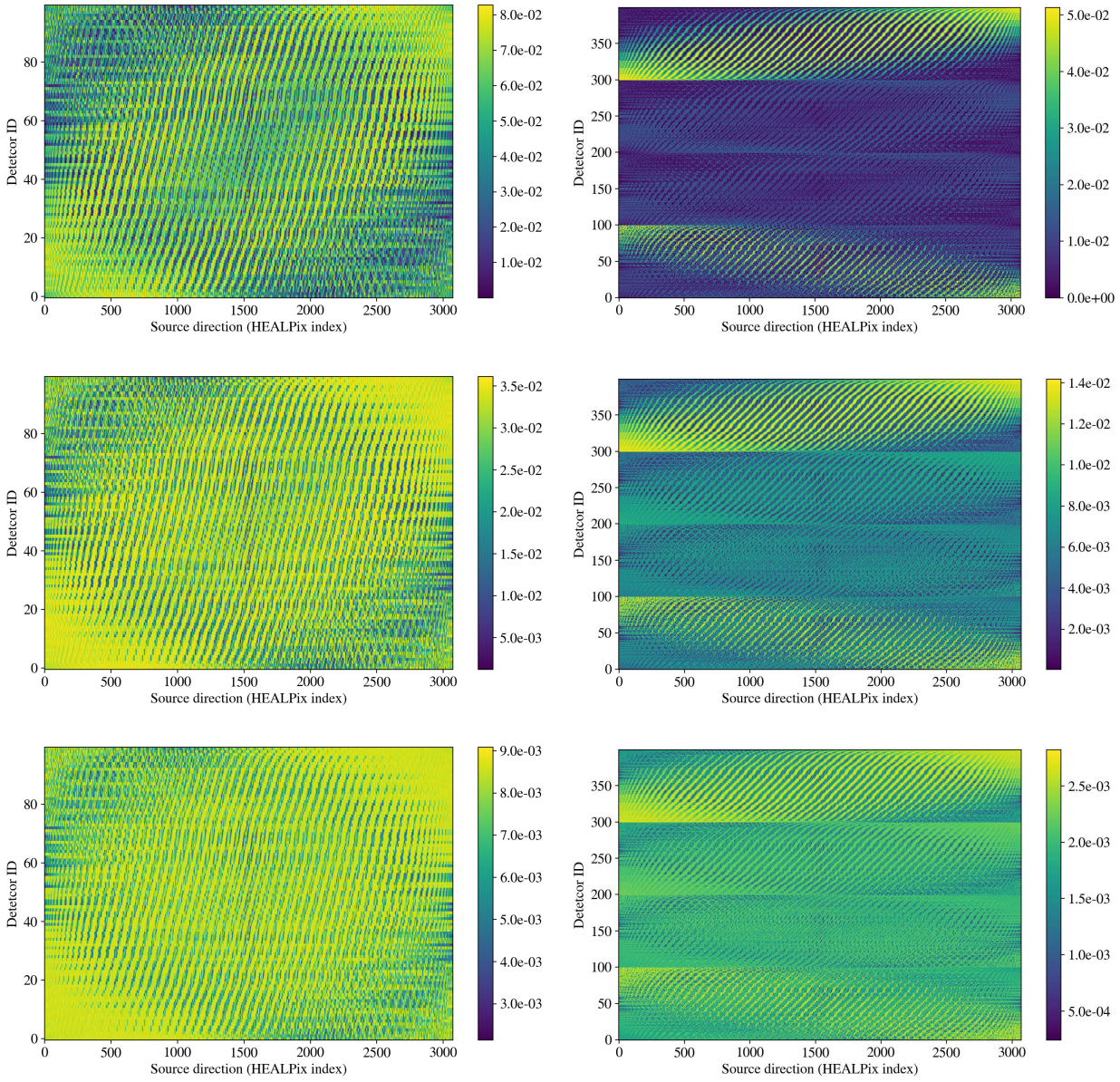

Figure 3.17. Far-field full-energy $N_{\text{side}} = 16$ directional system response of a random 100-detector PRISM configuration to 60 keV (top), 186 keV (middle) and 356 keV (bottom) gamma-rays, simulated with Geant4. The left column shows the response with no depth sensitivity and the right column with four DOI bins.

to simulate and store the path lengths of each ray through each detector, keeping track of the ordering of detector traversal. The path lengths can be used with the interaction cross-sections to calculate and accumulate the full-energy absorption probabilities in each detector for each ray, accounting for all occluding material that can absorb or

scatter the rays. The ray-tracing simulation can be run with a fully populated sphere and detectors selectively removed from the calculation to account for any mask configuration. This approach was used to determine the system response of the optimal PRISM configuration and used in the image reconstructions shown in Ch. 4.

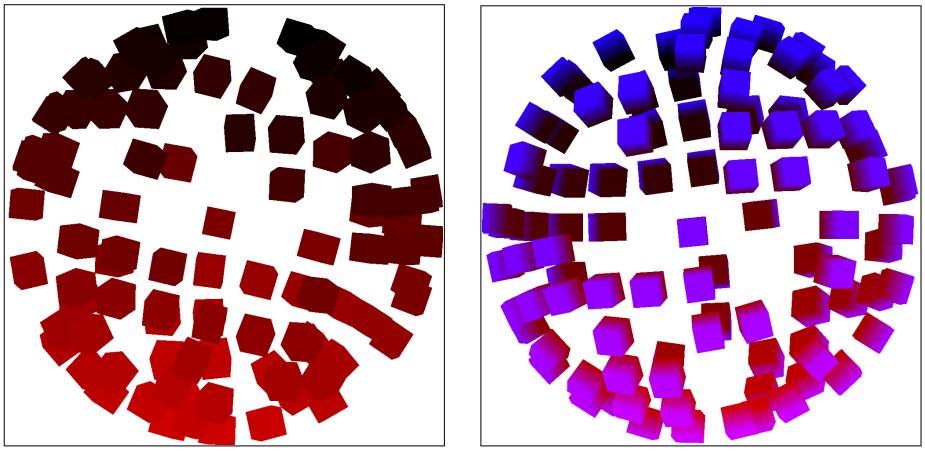

Figure 3.18. OpenGL display without (left) and with (right) DOI coloring.

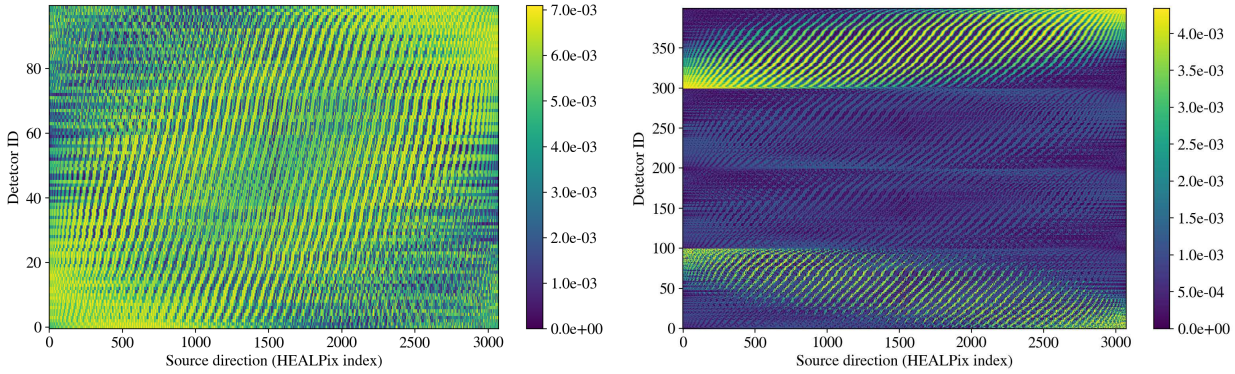

Figure 3.19. Zero-energy $N_{\text{side}} = 16$ (3072 directions) response of a random 100-detector PRISM configuration without (left) and with (right) DOI.

It is important to highlight the directional response of a single detector in the PRISM sphere (i.e., a row of the directional system matrix). Figure 3.20 shows the far-field zero-energy response of two detectors on opposite sides of the random 100-detector configuration, shown in a 4π Cartesian projection. The response is shown without any depth sensitivity (top row) and from only the inner (cathode side) depth bin (bottom row). The detector occlusions are clearly seen in the responses as checkered patterns of low intensity. A broad response is observed for the directions that illuminate the

crystals without any occlusions from other crystals. Four sections can be seen in this broad response with a depression in the middle, representing the increased surface area of a cube when viewed from the corners (high response) compared to the face (lower response). This response is almost completely removed in the inner depth bin as this bin is almost completely occluded by the rest of the crystal in these directions. In a list-mode formulation, these images are back-projected into the image space for an event in the detectors. As the reconstruction iterates through several counts in several detectors, each with their own response, the images combine to ultimately reveal the the source location.

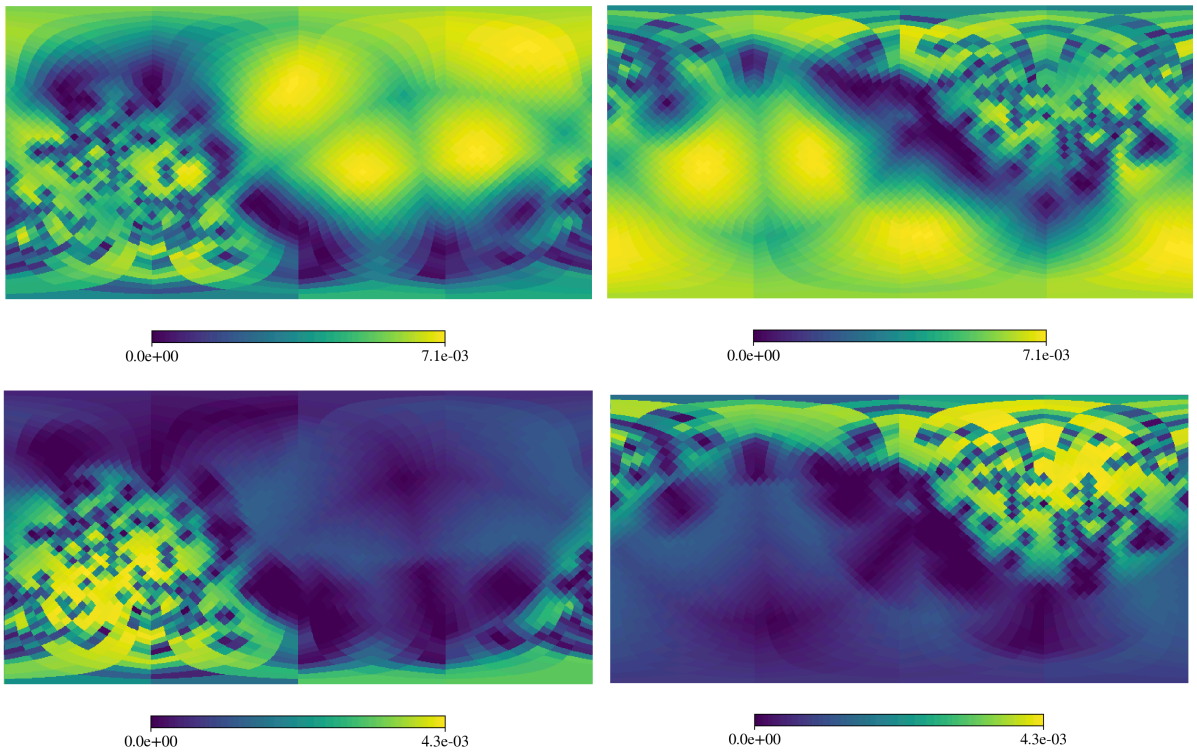

Figure 3.20. Zero-energy $N_{\text{side}} = 16$ response of two detectors on opposite sides of the random 100-detector configuration. (Top row) No depth sensitivity. (Bottom row) Inner (cathode side) depth bin.

3.6 Mask Pattern Optimization

An ideal gamma-ray imaging and search instrument would balance sensitivity (to efficiently detect weak sources) and imaging capability (to accurately localize sources). Note that energy spectroscopy and source identification capabilities are also important

in the source search problem when the source material is unknown. However, these capabilities are subject to the limitations in the detection material and the spectral analysis algorithms used for identification, and thus not the focus of this work. Active coded mask imaging relies on unique modulation patterns across the populated detectors to reconstruct the direction of potential sources. A fully populated (192 detectors) system would be highly sensitive (due to total detector mass), however will lack in unique modulation due to the regular grid spacing on the sphere. As detectors are removed, sensitivity is lost but unique modulations patterns arise. If too many detector are removed, however, unique modulation will also begin to degrade. Therefore care must be taken in determining the number and configuration (i.e., mask) of the detectors in the PRISM sphere.

In order to design an optimal mask pattern, a specific metric must first be chosen. There are many important performance characteristics of an omnidirectional imaging system, and the significance of each may vary depending on the application at hand. Therefore, in general, the choice of metric is often unclear. When designing passive planar coded apertures, a common metric is the maximization of the SNR of the reconstructed images. The URA [16] has been shown to optimize this metric [110] while also having constant sidelobes in its periodic autocorrelation function. The URA, however, is only applicable for a limited set of specific aperture sizes and detector geometries. Other mask types also exist, such as the non-redundant array (NRA), pseudo-noise array (PNP), and modified uniformly redundant array (MURA) [1], [15]. However, none of these masks have ever been designed for a spherical system. Therefore, an iterative approach similar to [111], [112] is taken here.

3.6.1 Approach

In this approach, many mask patterns (a single pattern is represented here with H) are simulated and evaluated (through the calculation of some metric or objective function, Q). The 192 available detector locations in PRISM present an enormous search space of possible mask configurations ($> 10^{30}$ after accounting for symmetries), therefore an exhaustive search approach is not feasible by any means. Moreover, the search space is highly non-convex, subjecting traditional gradient-based methods to local minimum/maximum convergence. Several solutions exist to this problem, with Simulated Annealing (SA) [113] being a well-known algorithm. In contrast to traditional convex algorithms that always reject steps that do not improve the objective function, these so-called heuristic global optimization algorithms tend to utilize some form of the Metropolis-Hastings algorithm [114], [115] that can accept steps in the optimization that do not improve the objective (the acceptance/rejection can be deterministic or probabilistic depending on the implementation) thus allowing the optimization to explore more space. SA takes this approach, using a probabilistic acceptance criterion based on the

"temperature" of the system. Furthermore, similar to annealing in metallurgy, this temperature is "cooled" over time, decreasing the chance of accepting a worse solution with increasing iterations. The problem with SA is that it can be quite sensitive to the choice of initial temperature and the cooling rate, or the "annealing schedule".

Dueck [116] proposed the heuristic Great Deluge (GD) algorithm as a single-parameter adaptation of SA and showed improvement with GD over SA for several optimization problems. The GD algorithm, for maximization, works in analogy to the tendency of a mountain climber to avoid being drowned in a flood, finding their way to the top of the mountain without getting their feet wet as the water level rises. For minimization, one can imagine a fish in a drought (decreasing water level), with the goal of staying underwater. In this sense, the solution at any given step is not compared to other solutions, but only to the water level, W. The water level increases or decreases deterministically based on the distance between the accepted solution, the water level, and a scaling parameter, d, between 0 and 1 meant to control the speed of convergence [111]. Here the optimization is formulated as a minimization, therefore W is decreased adaptively as:

$$W_{k+1} = W_k - d(W_k - Q_k)\,, \tag{3.1}$$

where W_k is the current water level, W_{k+1} is the water level of the next step, Q_k is the current metric value. A value of $d = 2.5 \times 10^{-2}$ was used in this work, as smaller values resulted in impractical runtimes.

The optimization begins with H equal to a random number of detectors in a random configuration on the sphere. The mask is perturbed (H^*) by toggling (on/off) a randomly sampled detector location. The metric Q is calculated before and after the perturbation and the GD criterion is used for acceptance. If accepted the old mask H is replaced with H^*, the water level is decreased according to Eq. 3.1, and the process repeats. If all 192 detector slots are toggled without lowering W, then two randomly selected detector slots are toggled. If the water level is lowered in the double toggle mode, the procedure reverts back to the single toggle mode. If the water level is not lowered over a predefined tolerance (i.e., 1000 iterations), the optimization is stopped and returns H as the optimal mask. The visual representation of the algorithm is given in Fig. 3.21.

To characterize a given mask, the ideal 2D imaging performance of far-field point sources at every pixel in the image space is analyzed. The directional system response was simulated using the zero-energy OpenGL approach and images were generated using sensitivity-corrected back-projection. The zero-energy approach was necessary to reduce computational time as tens of thousands of masks were evaluated. The OpenGL display utilized 9×10^4 pixels, as the addition of more pixels did not significantly change the computed system response ($< 1\%$ fractional change in the l^2-norm of the difference). With this many pixels, the system response can be computed in < 3 s.

For this work, the metric, Q, was designed to maximize sensitivity and localization performance, and force uniform performance over 4π. For a given far-field source in

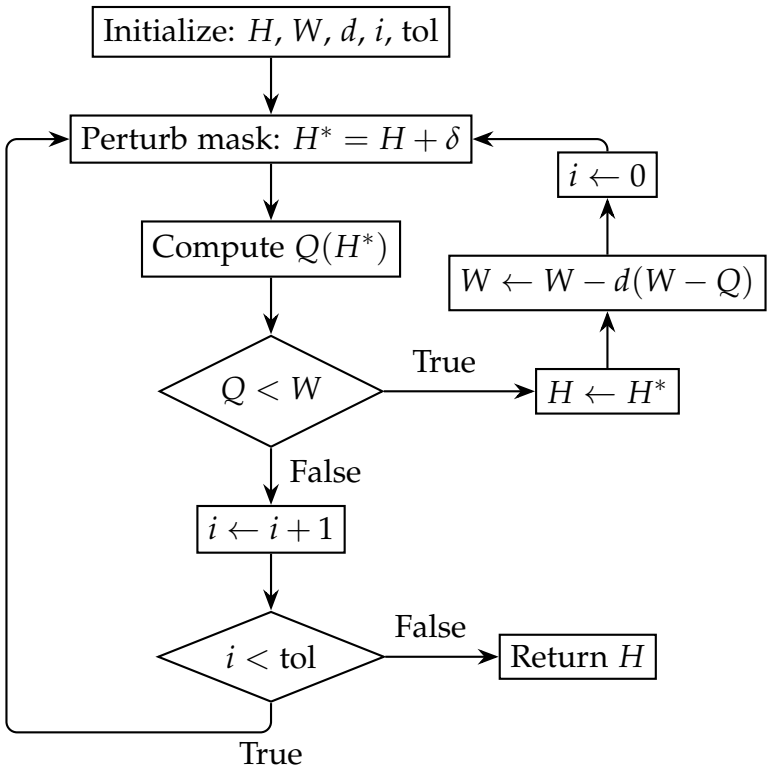

Figure 3.21. Flow chart of Great Deluge algorithm used for coded mask optimization.

direction (pixel) j, the sensitivity is defined as the detection probability in any detector, defined previously in Eq. 2.22. The localizing performance is described by a signal-to-blur ratio (SBR) of the sensitivity-corrected back-projection $\mathbf{w}^{\mathrm{BP}}$, given by

$$\mathrm{SBR}_j = \frac{\mathrm{w}_j^{\mathrm{BP}} - \mu(\mathbf{w}_{\bar{j}}^{\mathrm{BP}})}{\sigma(\mathbf{w}_{\bar{j}}^{\mathrm{BP}})}, \tag{3.2}$$

where $\mathrm{w}_j^{\mathrm{BP}}$ is the image intensity in the true pixel j, and $\mu(\mathbf{w}_{\bar{j}}^{\mathrm{BP}})$ and $\sigma(\mathbf{w}_{\bar{j}}^{\mathrm{BP}})$ are the mean and standard deviation of the image intensities in all pixels other than j, respectively. The blur in this treatment refers to the systematic geometric blurring inherent in the back-projection. Background noise and counting statistics were not included in the images.

The mean and variance of each component (sensitivity and SBR) are then computed over all source locations, j. The metric, Q, is defined by averaging the variances (σ^2) and the reciprocal of the means

$$Q = (1/4)\left[\mu(\boldsymbol{\varsigma})^{-1} + \sigma^2(\boldsymbol{\varsigma}) + \mu(\mathbf{SBR})^{-1} + \sigma^2(\mathbf{SBR})\right]. \tag{3.3}$$

In every iteration, each component of Eq. 3.3 is normalized by the first iteration value. Therefore the initial value of Q will be equal to 1. The overall minimization of Q will

maximize the mean sensitivity and SBR metrics and minimize the variation of the two metrics over 4π.

3.6.2 Results

The optimization procedure was performed six times with random initial detector configurations. Due to the large configuration space and the choice of d in Eq. 3.1, it is not expected that the results of the algorithm to necessarily be unique. In addition, it is not expected there to be a single mask configuration that significantly outperforms all others. In the several optimizations that were performed, the resulting masks differed slightly in design, however, were similar in both Q and the number of detectors (relative standard deviations of 1.5% and 2.5%, respectively).

Figure 3.22 displays the progression of Q as a function of iteration for a single random starting point and Fig. 3.23 shows the resulting optimized mask. For this particular run, the optimized mask contains 104 total detectors (out of the 192 total available locations). This is slightly larger than a 50% population, determined by [16] to be optimal for the URA. However, because there is a significant amount of open space even in the fully populated geometry due to packing cubes on a sphere and non-zero detector housing thickness, there is no reason to expect that 50% would necessarily be optimal in this case.

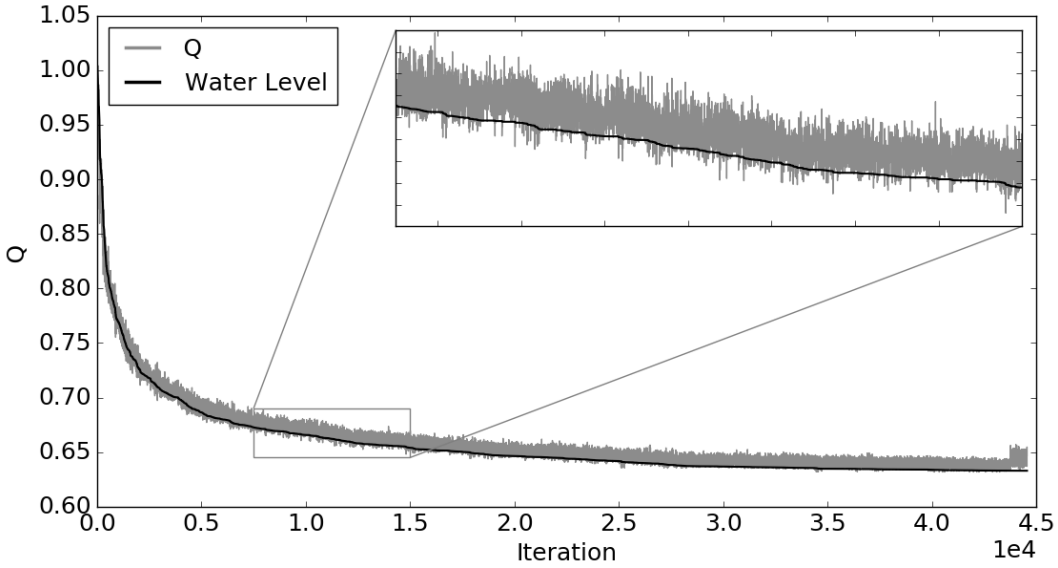

Figure 3.22. Minimization of Q using the Great Deluge optimization algorithm. At each iteration, the current mask configuration is retained if Q (gray) is below the water level (black).

Table 3.1 shows the individual components of Q between the starting and optimized masks. The optimized mask has a higher mean sensitivity (primarily due to the larger

number of detectors) and has much more uniform sensitivity over 4π. The optimized mask actually has a slightly lower mean SBR, but the uniformly of the SBR was greatly improved. It is expected that the behavior of these components can fluctuate because each were given equal weighting. Future work could be done to optimize the weightings of each component, depending on the application at hand.

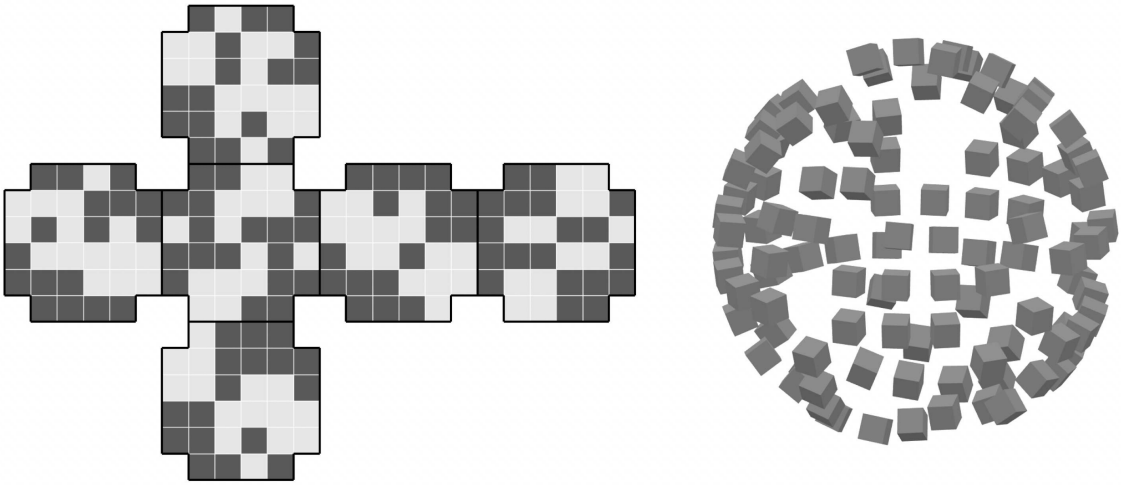

Figure 3.23. (Left) Optimized mask pattern shown with the modular faces of the spherical design unfolded into 2D. Filled in squares represent populated locations. (Right) Optimized mask shown with bare crystals in 3D. The mask contains 104 detectors out of the 192 available locations.

Table 3.1. Individual sensitivity and SBR metric values in the optimization metric, Q.

Mask	Detectors	$\mu(\boldsymbol{\varsigma})$	$\sigma^2(\boldsymbol{\varsigma})$	$\mu(\mathbf{SBR})$	$\sigma^2(\mathbf{SBR})$
Starting	84	0.366	2.9×10^{-4}	4.883	0.244
Optimized	104	0.431	8.0×10^{-5}	4.562	0.078

Chapter 4

Imaging

4.1 Omnidirectional Far-field 2D Imaging

The first task of assessing the performance PRISM as a gamma-ray imager was to demonstrate static far-field 2D imaging. The spherical arrangement of detectors is the first design to facilitate a 4π coded FOV for active coded mask imaging. To demonstrate this capability, four independent 10 min measurements were performed using PRISM-v0 with a 20 μCi ^{241}Am source (60 keV) at various directions in 4π with a 1 m standoff. See the top pane of Fig. 4.1 for a diagram detailing the source positions. The far-field approximation is appropriate at this distance given the detector size and expected angular resolution. Near-field effects will be discussed in Sec. 4.3. Image reconstruction was done with 50 iterations of ML-EM and the the results are shown in the bottom pane of Fig. 4.1, with the true source locations denoted with white cross-hairs. The intensity scaling has units of relative intensity as the system matrix used in the reconstruction was not quantitative. The image space was discretized into 3072 pixels according to the HEALPix scheme. The final reconstructed image was interpolated into a Cartesian grid projection with single degree discretization.

At the time of this measurement, the detector mask contained 93 detectors in a pseudo-random (not optimized) pattern. This is because the system was still being tested and developed at this point. Due to operation issues, some detectors exhibited a significant amount of low-energy noise and therefore the reconstruction only used data from 74 detectors.

The reconstruction was successful in accurately localizing the source in all four directions. This includes a source behind the system, where most of the detector electronics are placed. It is important to note that this measurement was performed with PRISM-v0, where most of the highly occluding material such as the battery are stored external to the detector enclosure. The PRISM-v1 system integrates both the onboard computer and Li battery in the back compartment, both of which will contribute to more attenua-

tion and scattering. Some non-uniformity is observed in the width and overall shape of the localized hot-spots for the four images. This measurement utilized a random mask pattern and thus a uniform imaging capability was not expected.

Given the hand-held design of PRISM and the development of SDF, the focus of this dissertation is on the free-moving 3D imaging capabilities of the system. Therefore, apart from the short discussion of near-field effects and DOI for 2D imaging (in Sec. 4.3 and Sec. 4.4), this work will primarily discuss 3D imaging.

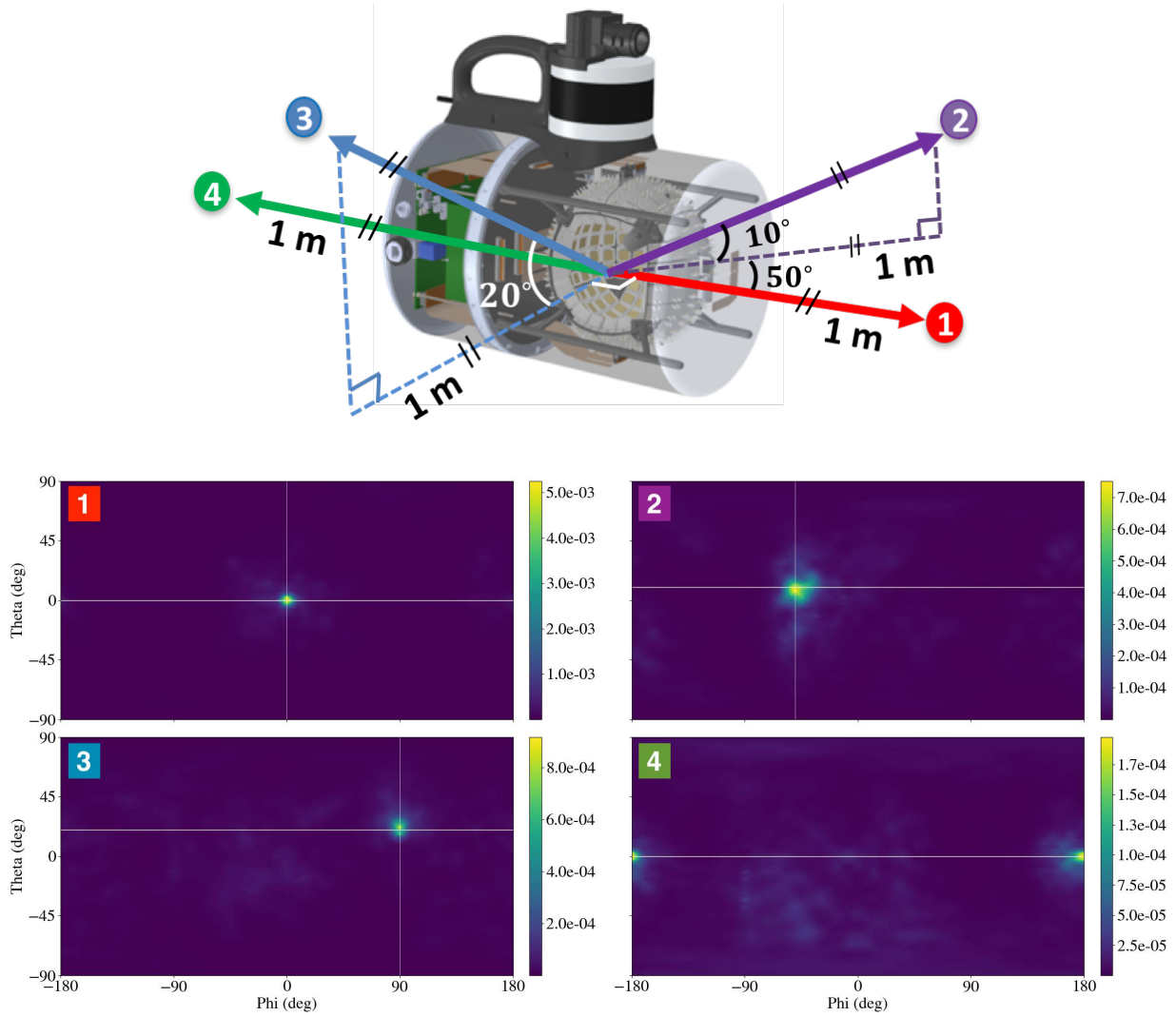

Figure 4.1. Static 2D images (50 iterations ML-EM) for 10 min measurements of 20 μCi ^{241}Am source placed at various directions in 4π with 1 m standoff. True source locations are denoted with white cross-hairs. Detector mask contained 93 detectors in a pseudo-random pattern (not optimized). Due to operation issues, the reconstruction only used data from 74 detectors.

4.2 Free-moving Far-field 3D Imaging

4.2.1 SDF Constraint and GPU Acceleration

In this section, free-moving 3D active coded mask imaging and the effects of the SDF constraint and GPU parallelization on both image quality and computational efficiency are demonstrated. An unshielded point-source localization scenario in a small indoor cluttered laboratory space is considered. A 20 μCi ^{241}Am source was placed in the corner of a large ($\sim$1 m^3) plastic case, approximately 25 cm from the top of the case. The PRISM-v1 system was walked around the lab, mapping the $\sim$70 m^2 space in less than 35 s (total of 320 poses). The 3D space surrounding the measurement path was uniformly discretized with 10 cm voxels, resulting in a total of 2×10^6 voxels and 5×10^4 occupied voxels (2.5% of the total). The list-mode gamma-ray data used in the reconstruction was limited to a 10 keV wide ROI around the ^{241}Am photopeak at 60 keV. A total of 1095 counts were collected in the ROI over the measurement from 92 detectors.

For comparison, the ML-EM reconstruction was performed on both the full and SDF-constrained image space and on both the CPU and integrated GPU. Ten iterations were used to balance image quality and overall reconstruction speed. Sample run-times for the sensitivity calculation and ML-EM iterations on the onboard CPU and integrated GPU are shown in Table 4.1. The full and constrained images are shown in Fig. 4.2 as 3D contour plots superimposed on a downsampled point cloud with top-down and isometric viewpoints. The color scales have units of relative intensity (arbitrary units) as the system matrix used for reconstruction is currently not quantitative.

The full voxelization reconstruction results in a bias towards the measurement path and a rather diffuse estimate in space. The SDF constrained image localizes the source with a higher degree of both accuracy (hotspot $<$ 10 cm from true location) and precision (tens of cm FWHM). In addition to image quality, the SDF constraint reduces the overall dimensionality of the reconstruction, leading to significant speed increases on both hardware. The GPU-based reconstruction shows a drastic increase in reconstruction speed over the CPU ($\sim$10$\times$), improving overall run-times to $\mathcal{O}$(s).

Table 4.1. Sensitivity and list-mode ML-EM iteration (itr) run-times on the PRISM-v1 single-board computer for the images in Fig. 4.2.

Hardware	Sensitivity (s)		MLEM itr (s)	
	Full	SDF	Full	SDF
Intel i7 dual-core 3.50 GHz (CPU)	550.8	14.2	49.0	1.1
Intel Iris Plus Graphics 650 (GPU)	47.3	1.3	7.5	0.2

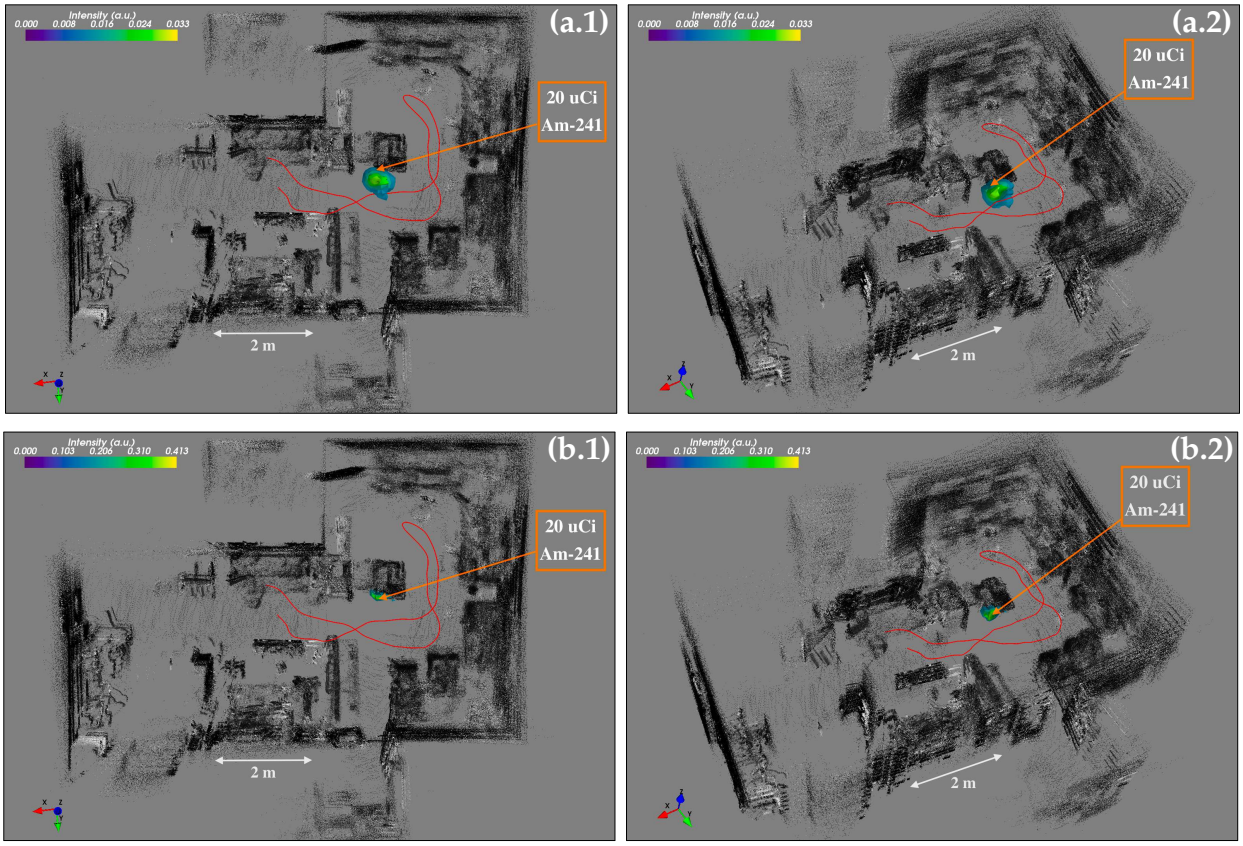

Figure 4.2. ML-EM reconstruction (10 iterations) of a 35 s free-moving measurement (path shown in red) around a ^{241}Am source (marked with an arrow) in a small indoor cluttered lab space ($\sim$70 m^2). (a) Full voxelized model [(1) top-down and (2) isometric viewpoints] and (b) SDF-constrained model [(1) top-down, (2) isometric], superimposed on the 3D point cloud colorized with LiDAR return intensities. Measurement parameters: 10 cm voxel size, 5×10^4 occupied voxels (2.5% of total), 320 poses, 92 detectors, and 1095 counts in the photopeak ROI (59 ± 5 keV).

4.2.2 OS-EM Acceleration

An SDF-constrained GPU-accelerated 3D OS-EM reconstruction algorithm (both list-mode and bin-mode) was implemented using the OpenCL framework. In the bin-mode formulation, sets are defined as random groups of detectors. The free-moving nature of the measurements facilitates balanced sets as counts are equally distribution across detectors as the system is moved through the environment and views sources from various perspectives. In the list-mode formulation, sets are defined as random groups of events. The shuffling of events ensures that each set will have a similar number of counts that contribute to the overall source localization. OS-EM is known to have convergence issues in the presence of noise and can be sensitive to number of sets and iterations. With appropriate statistics and small number of sets and iterations, however, OS-EM can

result in accelerated reconstructions (i.e., similar images to ML-EM in less time).

This method is explored using the same measurement scenario from the previous section (see the caption of Fig. 4.2). First, a comparison is made between ML-EM and OS-EM operating in bin-mode. Figure 4.3 shows the image results (top-down projection contour plot on a downsampled point cloud) and total reconstruction times using various numbers of sets and iterations. Recall that a single OS-EM iteration includes the iterations through all subsets. The iterations through the subsets are referred to here as inner iterations and the OS-EM iteration is referred to as an outer iteration. Note that OS-EM with one set is equivalent to ML-EM (i.e., each outer iteration contains a single inner iteration). It is observed that the OS-EM reconstructions can achieve approximately the same degree of convergence in less amount of time. Specifically, the reconstruction with four subsets (groups of 23 detectors) results in a more precise localization with 5 iterations in roughly half the time of ML-EM with 10 iterations (middle image). The reconstruction with 23 subsets (groups of 4 detectors) localizes the source with only one iteration, with a speed increase of $\sim 6\times$ (bottom left image).

Given the measurement parameters (i.e., number of poses, detectors, counts), it is more efficient to run the reconstruction in list-mode. Figure 4.4 again shows the image results and total reconstruction times using various numbers of sets and iterations, though this time operated in list-mode. The first observation is that list-mode ML-EM significantly outperforms its bin-mode counterpart in terms of reconstruction speed (the image results at any given iteration are equivalent). However, the OS-EM reconstruction does not show an improvement in reconstruction speed as the number of subsets is increased. This is because the GPU is capable of efficiently distributing the problem over the low number of events ($\sim$1000), particularly in the forward-projection step, even in the one subset case. In other words, the forward-projection step requires the same amount of time, regardless of the number of subsets. Therefore adding more subsets increases the outer iteration time. While fewer iterations are required for similar convergence to 10 iterations of ML-EM, the increased forward-projection computation ultimately results in a slower overall reconstruction. This effect will vanish in the case of large numbers of events, where subsetting becomes favorable for the GPU implementation. In this case, a significant speed increase is expected with OS-EM. Finally, convergence issues are indeed observed in the list-mode OS-EM reconstruction when using a larger number of subsets (bottom row of images).

In some cases, OS-EM offers a substantial reconstruction acceleration over a traditional ML-EM. However, in its current form, the sensitivity to the choice of the number of sets and iterations limits its use in radiological search scenarios where *a priori* information that could aid in the choice of subsets is limited. Moreover, in count-limited scenarios where list-mode reconstruction is favorable, no significant gain is seen with OS-EM due to the GPU implementation. In this case, the GPU-accelerated list-mode ML-EM is the best solution in terms of both image quality and reconstruction speed.

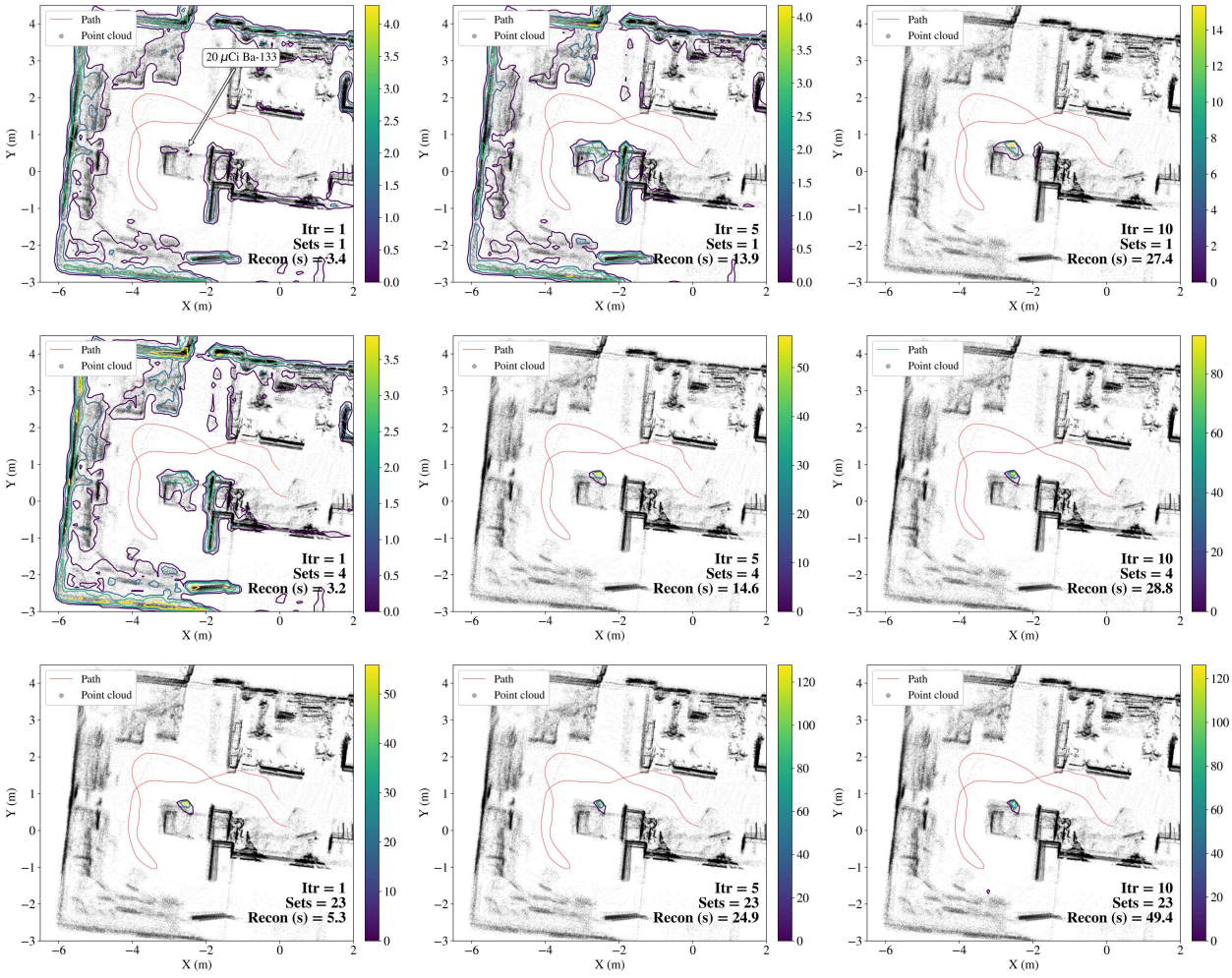

Figure 4.3. SDF-constrained GPU-accelerated bin-mode OS-EM image reconstructions (top-down projection contour plot on a downsampled point cloud) and sample computation times for the measurement initially presented in Fig. 4.2. Number of sets are varied across rows: 1 (top), 4 (middle), and 23 (bottom). Number of iterations are varied across columns: 1 (left), 5 (middle), and 10 (right). The plots along the anti-diagonal (top-right to bottom-left) highlight the ability of OS-EM to increase image quality and reduce computation time with increasing number of sets.

In the future, relaxation methods such as decreasing the number of subsets with each outer iteration could alleviate convergence issues, though the choice of initial conditions and relaxation rates are additional degrees of freedom that may not be well defined in unknown environments. Subsets across the larger reconstruction dimensions, such as detector poses or image voxels, could also be considered to better leverage the parallelization capabilities of GPUs.

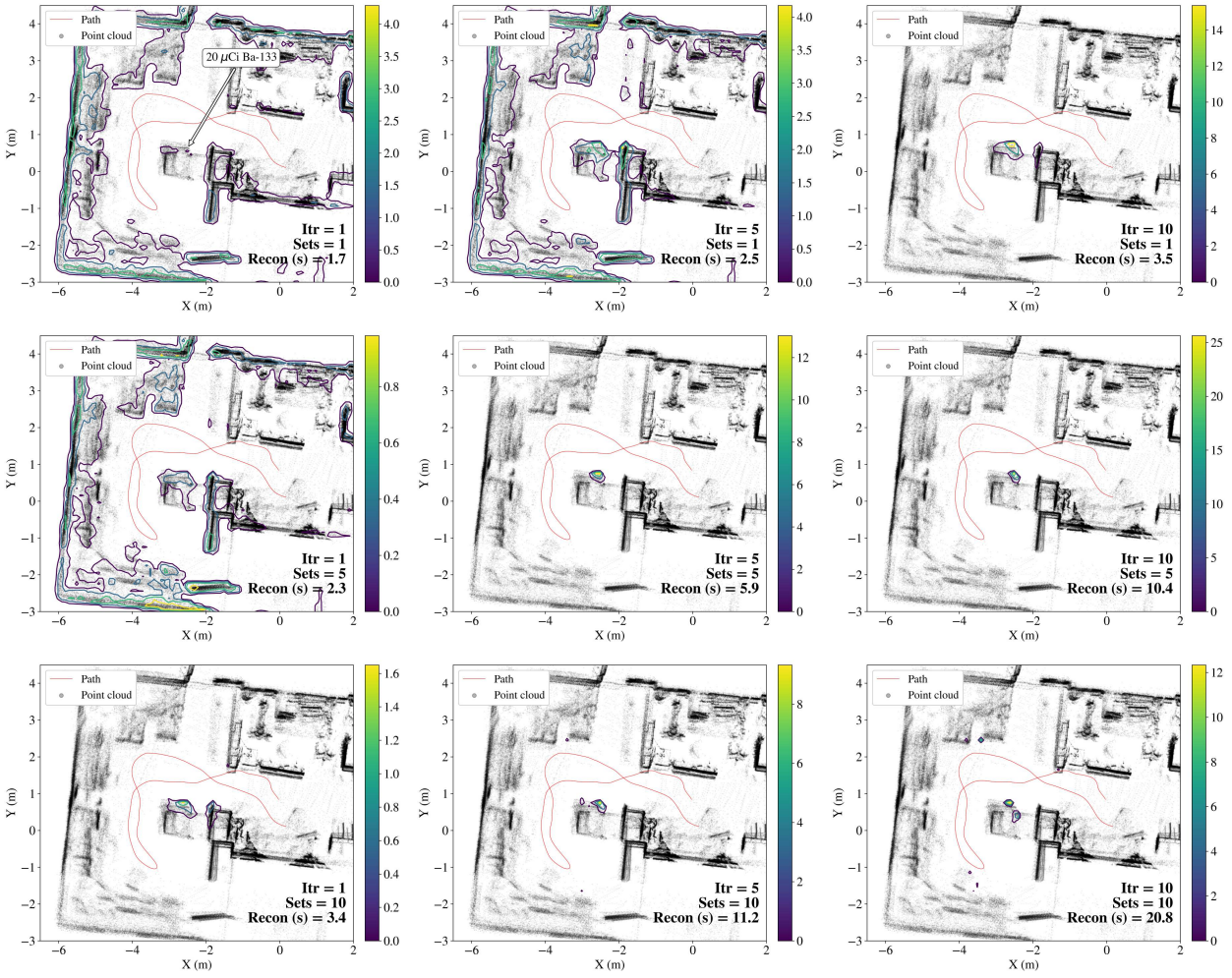

Figure 4.4. SDF-constrained GPU-accelerated list-mode OS-EM image reconstructions and sample computation times for the measurement in Figs. 4.2 and 4.3. Number of sets are varied across rows: 1 (top), 5 (middle), and 10 (bottom). Number of iterations are varied across columns: 1 (left), 5 (middle), and 10 (right). Sub-setting the space did not improve image quality or computational speed in this case.

4.2.3 Real-time Online Reconstruction

The results shown in the previous sections were for data collected over the entire measurement. However, Table 4.1 shows that the SDF constrained GPU-based imaging approach can facilitate online reconstruction and return images to the user in real-time during the measurement. In this section, the real-time online 3D imaging capability is demonstrated in an indoor office space, consisting of two large rooms connected by a hallway (total area of $\sim$240 m^2). A 20 μCi ^{133}Ba source and a 20 μCi ^{137}Cs source were

placed on top of filing cabinets on opposite sides of one room. The PRISM-v1 system was walked around both rooms in under 2 min. A 10 keV wide ROI was placed around 356 keV to localize the ^{133}Ba source and to demonstrate the coded mask imaging capability at higher energies. The 662 keV line from ^{137}Cs is better suited for high-energy imaging modalities (i.e., Compton imaging).

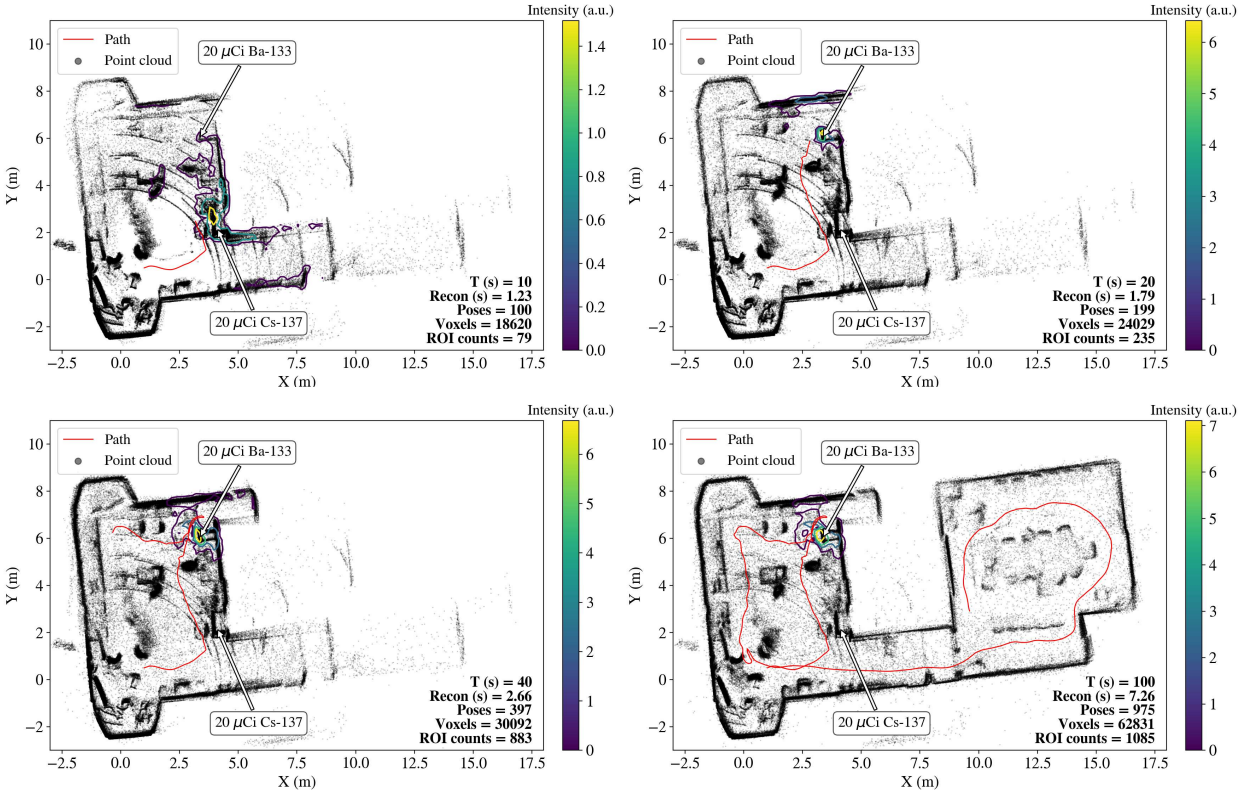

Figure 4.5. Time sequence of the top-down projection SDF-constrained ML-EM reconstruction (10 iterations) during a free-moving measurement (shown with a red line) in a small indoor space ($\sim$240 m^2). Two sources (^{133}Ba and ^{137}Cs) were placed in the scene and are marked with arrows. Reconstruction was performed on the high-energy ^{133}Ba photopeak (356 $\pm$ 5 keV). The source was correctly localized within $\sim$20 s.

Figure 4.5 shows a series of top-down projection images (10 iterations ML-EM) including the point cloud and pose estimates during the course of the measurement, each with a timestamp and additional reconstruction parameters such as the number of poses, occupied voxels, counts, and the reconstruction time (sensitivity and all iterations) for the state of the measurement up to that point. The image reconstruction is always performed in 3D, and 2D projection is done only for visualization. Both the scene and the gamma-ray reconstruction update over time as new space is explored. The system first approaches the ^{137}Cs source and produces a localization estimate around its position,

likely due to the Compton continuum of the 662 keV gamma-ray in the 356 keV ROI. As the system gets closer to the ^{133}Ba source, however, the image quickly corrects and localizes the ^{133}Ba source. The solution converges in < 40 s and remains roughly static as the system continues to explore new space.

Figure 4.6 shows multiple views of the final 3D image, visualized as a colorized point cloud (processed and rendered offline). Points from the ceiling and floor have been removed for clarity. Several blurry tracks can be seen in the point cloud, particularly in the room with the sources. The points arise from people moving through the scene during the measurement. Techniques such as those in [75] could be explored to remove transient points in the point cloud.

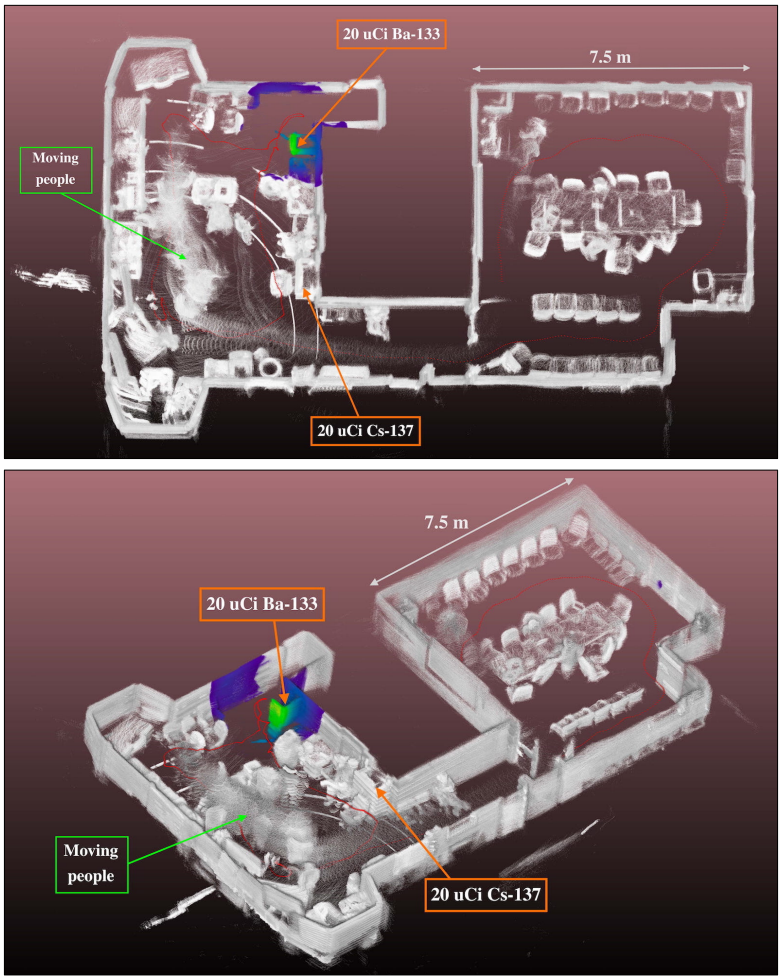

Figure 4.6. Final 3D image from Fig. 4.5, shown as a colorized point cloud (visualization processed offline). The ^{133}Ba source is correctly localized to the top of the cabinet in the back corner of the office. Point cloud tracks from people moving during the measurement can be seen in the source room. Points from the ceiling and floor have been removed for clarity.

4.2.4 Wide-area Outdoor Localization

In contrast to structured light sensors based on infrared light (e.g., Kinect), the use of LiDAR facilitates SLAM in wide-area outdoor settings. To demonstrate the performance of the 3D active coded mask imaging modality in this setting, PRISM-v1 was walked toward and around a single vehicle parked among several vehicles in a large open field. A strong plutonium surrogate [194 μCi ^{252}Cf, 530 μCi ^{133}Ba, 35 μCi ^{137}Cs] was placed in the open trunk. An area of $\sim$2,500 m^2 was mapped in under 1.3 min. The SDF-GPU ML-EM image reconstruction (10 iterations) was run on a 10 keV ROI around the low-energy 81 keV ^{133}Ba photopeak. Figure 4.7 shows the results of the reconstruction as a colorized point cloud from several viewpoints. The source is correctly localized to the trunk of the car, with a hot-spot accuracy on the order of tens of cm. Gamma-ray intensity is also seen on the open trunk door and the ground beneath the source and likely represents down-scattering from the primary source. In addition to 3D source localization in large outdoor settings, the results also highlight the benefit of LiDAR to produce dense, high-resolution, context-rich point clouds.

Another outdoor wide-area source search scenario measurement was performed at a site consisting of a collection of small and large house-like structures made of various materials (e.g., clay, brick) as well as a few steel L-shaped cargo containers. A 5 mCi ^{133}Ba source was placed in one of the cargo containers and PRISM-v1 was walked around the $\sim$4,500 m^2 site in less than 2.5 min. The SDF-GPU ML-EM image reconstruction (10 iterations) was run on a 10 keV ROI around the high-energy 356 keV ^{133}Ba photopeak, again to demonstrate the energy range of active coded mask imaging. Figure 4.8 shows the 3D reconstruction results as a colorized point cloud from several viewpoints. The gamma-ray hotspot was accurately localized to the correct cargo container among > 10 structures in the scene.

Since the SDF constraint currently limits the reconstruction to the point cloud (i.e., surfaces), voxels inside the containers are labelled as free and not included in the reconstruction. The source was placed inside of the container and therefore the reconstruction only shows the surface gamma-ray emission profile of the container. The implementation of a tri-state occupancy model (occupied/free/unknown) could be used to capture and reconstruct on voxels inside of closed volumes or unexplored spaces (e.g., behind the perimeter wall). Ray-casting techniques such as those used in [76] could be used for tri-state occupancy.

Finally, a laser-scan mismatch can be seen in Fig 4.8 (structures have multiple front walls), highlighting the fact that the gamma-ray reconstruction is currently limited by errors in the point cloud and pose uncertainties. SLAM algorithms are expected to improve with developments in autonomous vehicles, however detailed characterizations of SLAM uncertainties will be necessary to understand the impact on free-moving gamma-ray imaging and SDF.

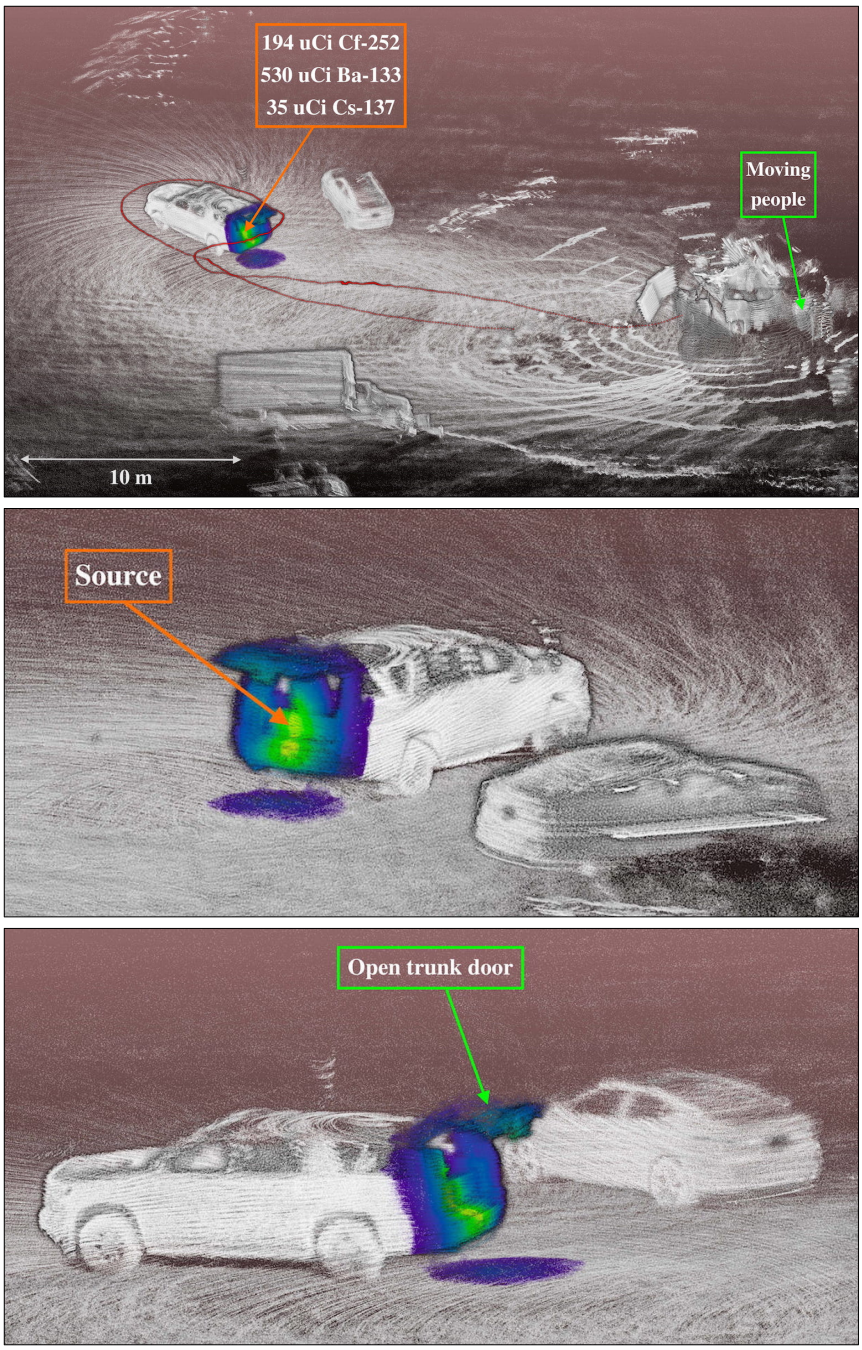

Figure 4.7. Final ML-EM reconstruction (10 iterations) from a 1.3 min measurement (path shown by the red line) around a vehicle containing a strong plutonium surrogate source in a wide-area outdoor setting (∼2,500 m^2), shown as a colorized point cloud (visualization processed and rendered offline). Measurement parameters: 15 cm voxel size, 7×10^4 occupied voxels (2.9% of total), 755 poses, 91 detectors, 4542 photopeak counts (81 ± 5 keV), 10 s reconstruction time.

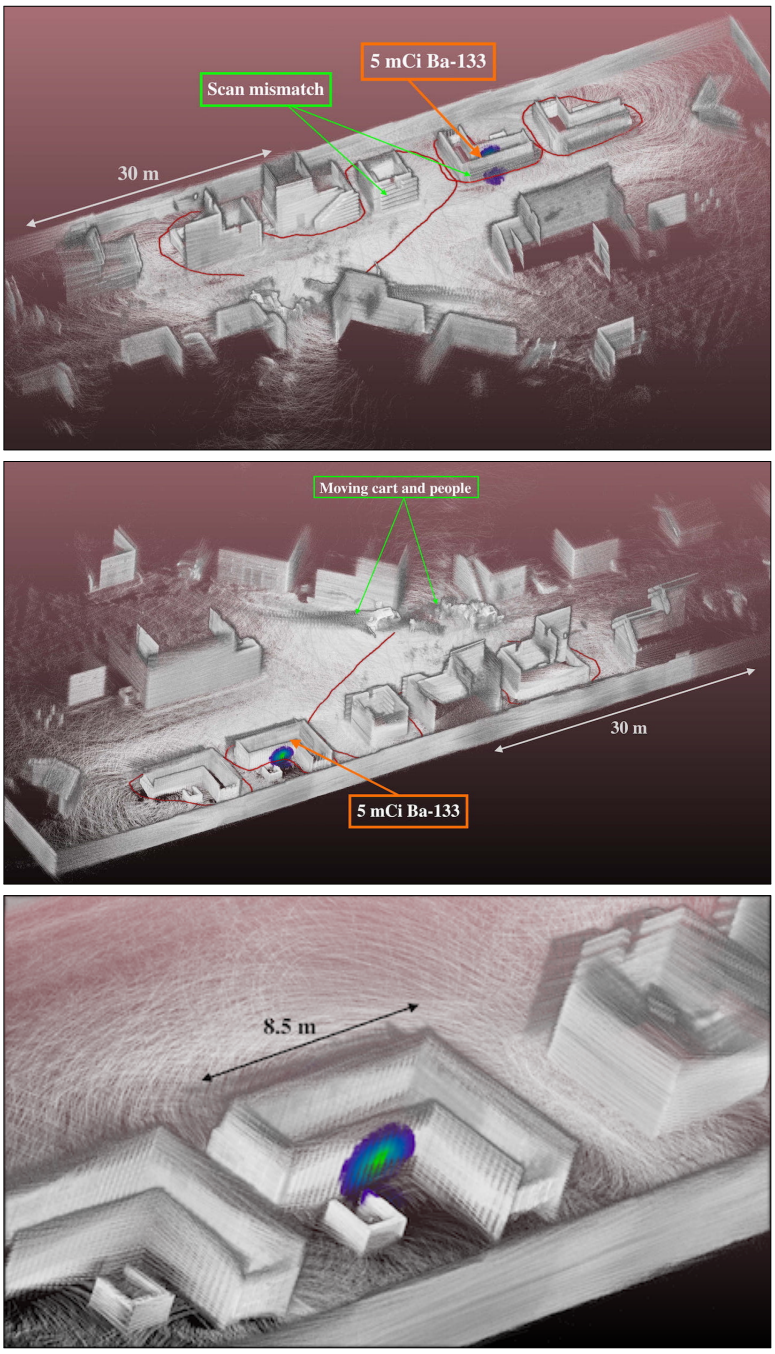

Figure 4.8. Final ML-EM reconstruction (10 iterations) from a < 2.5 min measurement (path in red) in a wide-area outdoor setting ($\sim$4,500 m^2) containing several house-like structures and steel L-shaped cargo containers. A ^{133}Ba source was placed inside one of the containers. Measurement parameters: 30 cm voxel size, 7×10^4 occupied voxels (7.3% of total), 1400 poses, 88 detectors, 9566 photopeak counts (356 ± 5 keV), 20 s reconstruction time.

4.2.5 Multi-source Reconstruction

Real-time imaging also facilitates imaging multiple energies (i.e., multiple isotopes) during a single measurement. A list of isotopes and associated gamma-ray energies could be defined before the measurement or during the measurement with a real-time spectroscopic detection algorithm. The reconstruction can then be run on each energy in a serial fashion and the images superimposed onto a single display.

To demonstrate this capability, a multi-source scenario consisting of ^{241}Am and ^{133}Ba sources (both 100 μCi) placed in close proximity in a small laboratory space is considered (see Fig. 4.9). PRISM-v1 was walked around the space in $<$ 38 s (370 poses) with 91 operational detectors. The image space was voxelized with 10 cm voxels (2.4×10^6 total voxels and 5.2×10^4 occupied voxels). A total of 3945 and 4114 events were recorded in the 10 keV wide ^{241}Am and ^{133}Ba ROIs, respectively. The reconstructions are shown as superimposed 3D contour plots in Fig. 4.10 and separate colorized point clouds in Fig. 4.11. The computational time of each reconstruction (10 ML-EM iterations) was $<$ 6 s.

Figure 4.9. Multi-source scenario setup. Placement of ^{241}Am and ^{133}Ba check sources indicated with arrows.

Both sources are accurately localized to the opposite corners of the large box (hotspots within $\mathcal{O}$(cm) to the true locations). The additional blur in the ^{241}Am image is most likely due to down-scatter (both in the detector and environment) from the higher energy gamma-rays from the ^{133}Ba source and ROI overlap with the 81 keV photopeak. The real-time performance of multi-source reconstruction will slow with more sources and larger

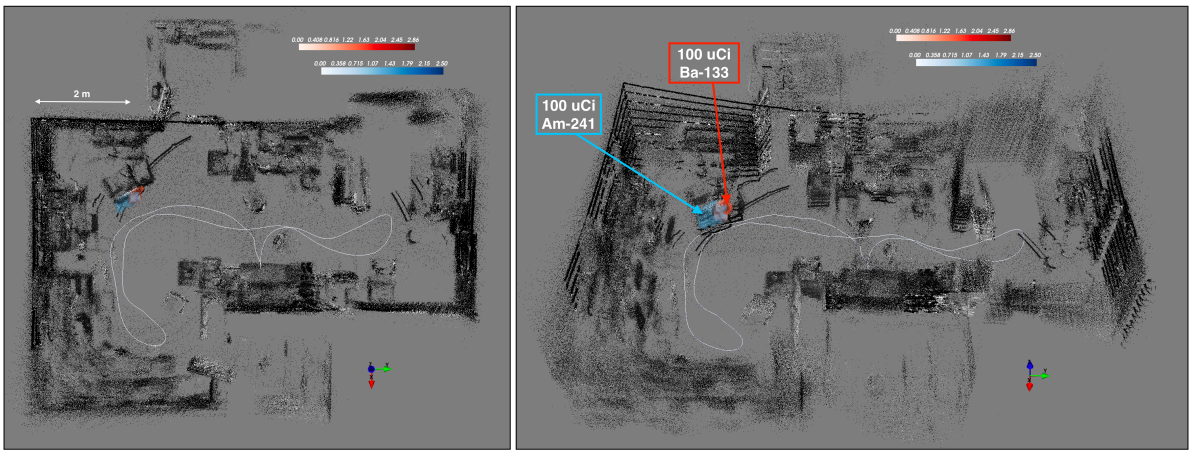

Figure 4.10. SDF-constrained ML-EM reconstructions (10 iterations) for a free-moving measurement of the source setup in Fig. 4.9 using 10 keV ROIs around the ^{241}Am and ^{133}Ba photopeaks at 60 and 81 keV, respectively. Measurement path is shown in white and true source locations are shown with arrows. Measurement parameters are given in the text.

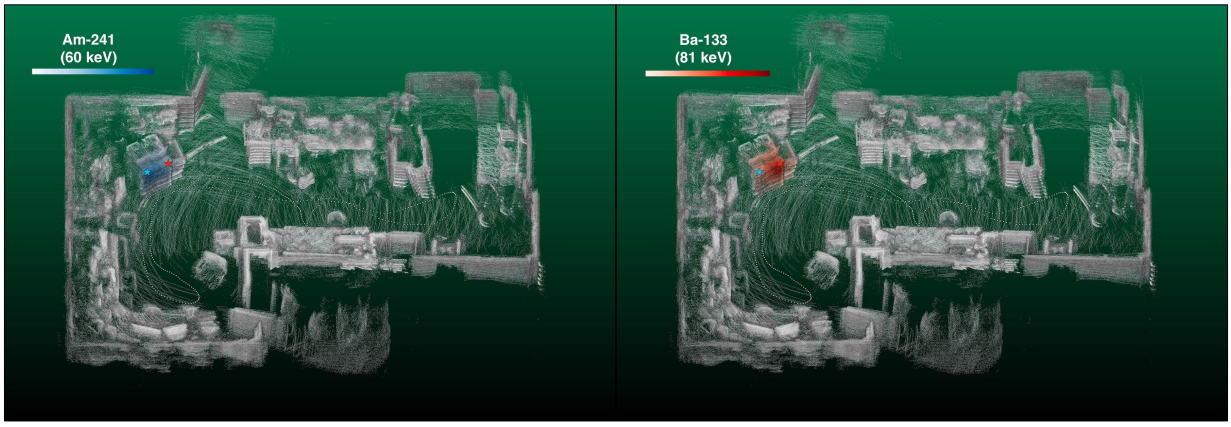

Figure 4.11. Images in Fig. 4.10 visualized as separate colorized point clouds. True source locations are shown with colored stars (^{241}Am blue and ^{133}Ba red).

reconstruction spaces, due computational demand from each individual reconstruction. In addition to increasing the voxel size over time or reducing the number of iterations, spectroscopic detection algorithms could be used to limit the spatial and temporal extent of each individual reconstruction (i.e., subsets of voxels and poses).

A more interesting approach would be to implement an "energy imaging" algorithm in which every event is reconstructed into both the 3D position space and the 1D emission energy space. Images could be decomposed into a 3D spatial image at any energy (e.g., 60 keV) or a full energy spectrum in any voxel. This concept has been explored previously at higher energies with Compton cameras [11], [117], [118], combining 2D spatial

reconstruction techniques of Compton imaging with spectral-deconvolution algorithms. PRISM and free-moving SDF methods could be used in the future to demonstrate energy imaging in all three spatial dimensions and lower energies.

One unique application of this capability, relevant for nuclear security and safeguards, would be 3D uranium enrichment imaging, in which the reconstructed spectrum in every voxel could be converted to an uranium enrichment estimate using traditional methods (i.e., photopeak ratios) [119], [120]. There exists issues of course with self-shielding, infinite thickness, and environment attenuation and scattering, though even coarse (depleted, natural, low enriched, highly enriched) and surface estimates would be useful in certain applications. For example, UF_6 is a common material in enrichment and reprocessing facilities and is typically kept in liquid or solid form in large cylinders for shipping and storage. Many of these cylinders are stacked together to create a 3D topology of UF_6. One could imagine a free-moving measurement through one of these facilities, creating a 3D map of the cylinders, each with an enrichment tag. Such a map could be used for material accountability or to aid in safety and safeguards inspections.

4.3 Near-field Effects

Hand-held imagers are efficient search tools as they can overcome the inverse square law by moving much closer to potential sources than traditional static imagers. However, as the system is brought close to the source, near-field effects can begin to dominate and, if not accounted for in the reconstruction, can produce artifacts and additional blurring. By properly characterizing and utilizing the near-field effects, the artifacts can be removed and spatial resolution can be increased. Near-field effects also facilitate high-resolution static 3D coded aperture imaging and has seen applications in medical and small animal imaging [5], [7], [121]–[123] as well as for position sensitivity in bulk scintillators [124].

4.3.1 Response Modeling

Traditionally, the system is small enough and the source distance is large enough that a far-field (point-detector) approximation can be made. Thus a 2D directional response is sufficient for reconstruction. In previous sections, PRISM has used 3072 directions defined on S^2 by the HEALPix discretization scheme and the response was generated either by Geant4 simulations, ray-tracing or through a fast zero-energy (i.e., infinite attenuation and no scattering) graphics-based simulation in OpenGL [41].

As the source is brought closer to the system, magnification and solid angle effects due to finite detector size will dominate and the far-field approximation breaks down. This effect can be captured in Geant4 by simply changing the incident photons from a parallel beam to a cone beam originating from the source location in 3D. To simulate

enough source positions in $\mathbb{R}^3$ to sufficiently sample the 3D space would be computationally expensive with Geant4, requiring many CPU-hr. Therefore the fast zero-energy graphics-based simulation was used here to sample the near-field 3D system response. While some physics is not included, the zero-energy approximation is reasonable at energies < 100 keV, as discussed in previous sections. To include near-field effects, the simulation was updated from an orthographic projection to a perspective projection. The reader is referred to [125] for diagrams of these two projections (specifically Figs. 5.13 and 5.14). Figure 4.12 shows the perspective projection of a randomly populated PRISM system at various distances from the center of the sphere. As a result of the perspective projection, a relative magnification/minimization is observed between the near-side and far-side detectors. A global correction was applied in post-processing to account for the minimized far-side detectors and normalize the responses (using the solid angle subtended by the un-occluded far-side detectors as a function of distance).

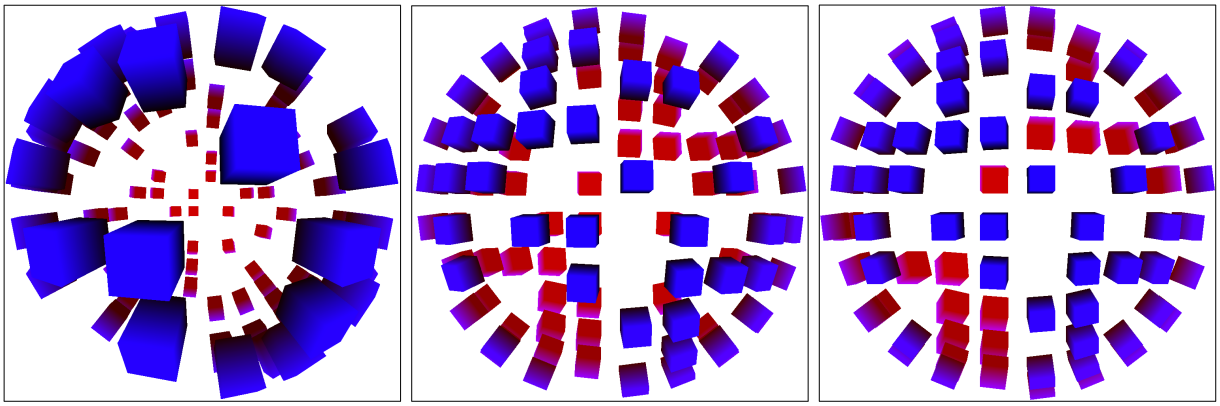

Figure 4.12. OpenGL perspective projection of a randomly populated PRISM crystal array rendered from distances of 10 cm (left), 50 cm (middle), and infinity (right).

Figure 4.13 shows example 4π responses (Cartesian projection of HEALPix) of a single detector in a random 93-detector configuration at various source distances. As the source is brought closer to the un-occluded side of the detector (broad distribution shown in yellow) the directional response increases in intensity (due to the inverse square law and larger solid angle) and decreases in extent (due to the occlusion of adjacent detectors). Figure 4.14 shows the response of all 93 detectors from a source at various distances along a single direction (i.e., the shadowgram). In addition to the overall scaling due to the inverse square law, the distinct differences in the response is what facilitates static near-field 3D imaging.

To demonstrate the effects of using a far-field response to reconstruct a near-field source, a simple 2D imaging scenario is considered. The zero-energy simulation was used to generate the 4π system matrix at a distance of 20 cm from the center of the PRISM sphere ($\sim$12 cm from the detector enclosure). A column of the system matrix

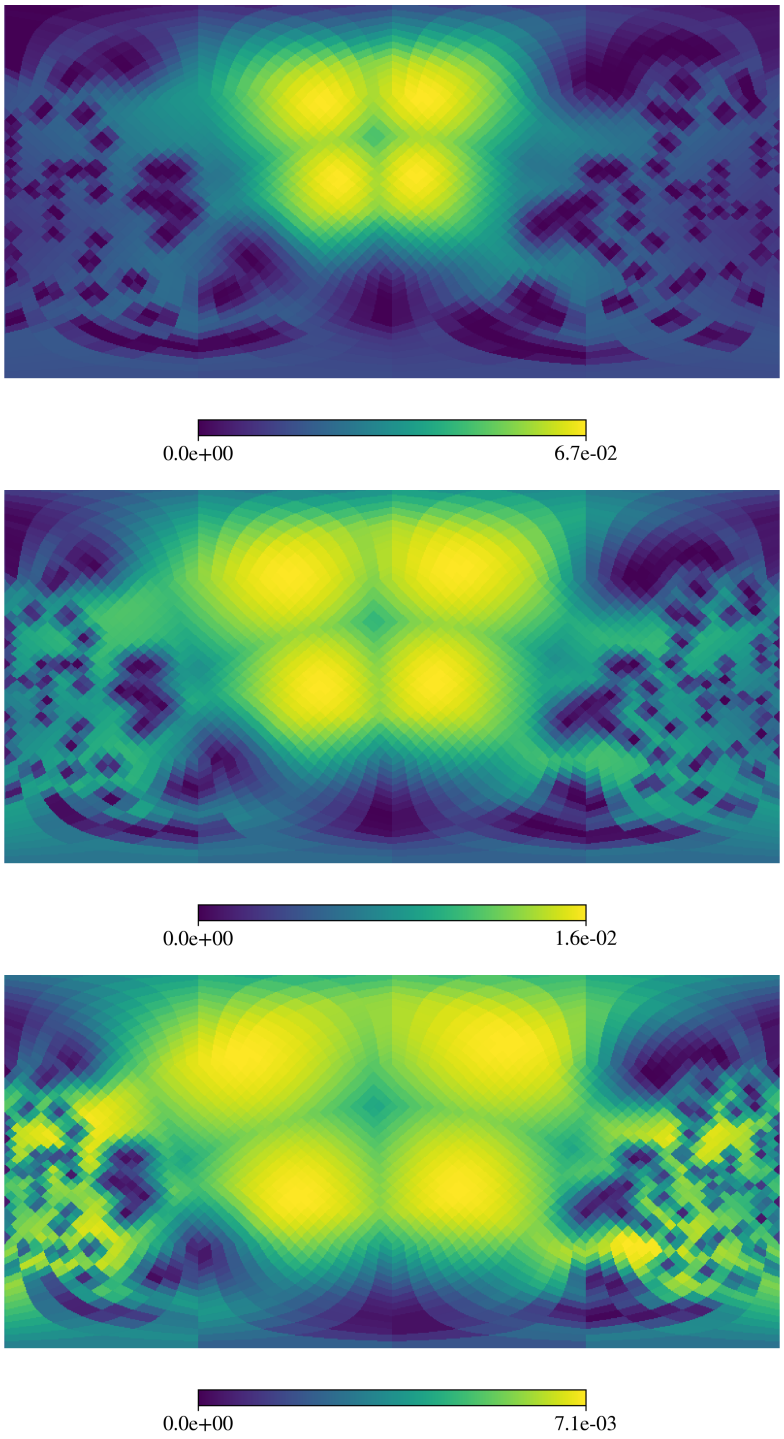

Figure 4.13. Single detector 4π zero-energy directional response at source distances of 20 cm (top), 50 cm (middle), and infinity (bottom).

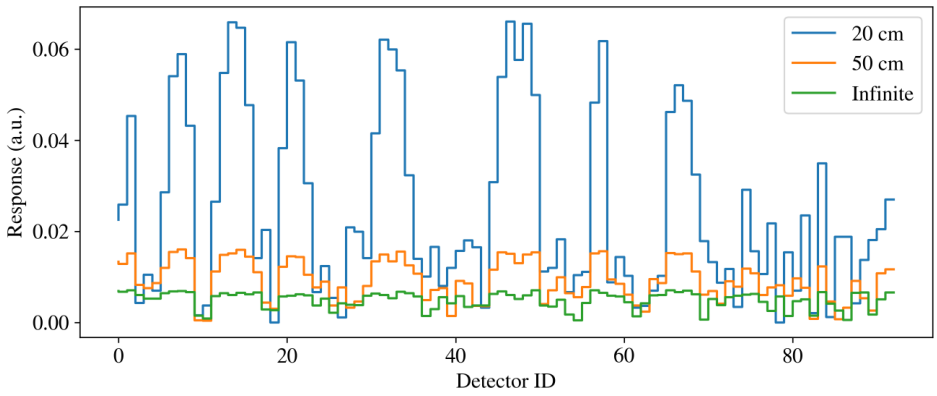

Figure 4.14. Zero-energy detector responses from a point-source at various distances along a single direction.

(single direction) was used as the input signal to image reconstruction. Figure 4.15 shows the 2D ML-EM image reconstruction (20 iterations) in 4π using the far-field and 20 cm near-field directional responses. Blurring and artifacts are observed with the far-field approximation, but are removed when near-field effects are properly accounted for.

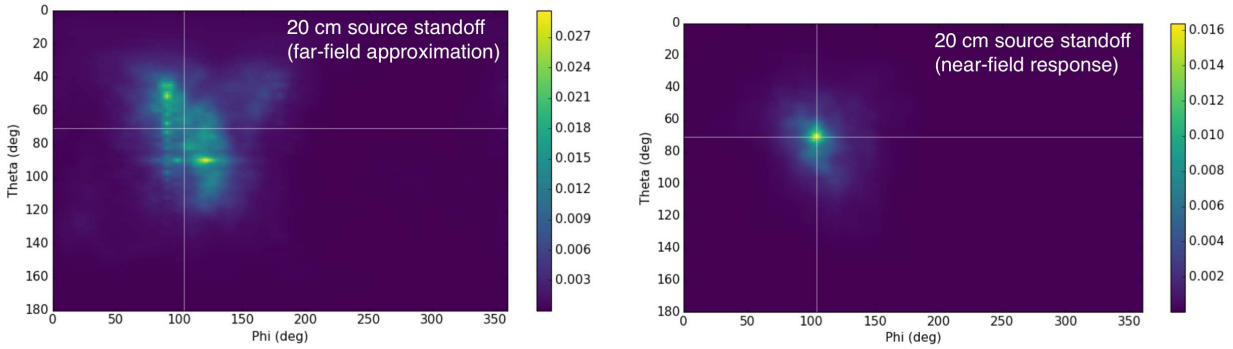

Figure 4.15. ML-EM image reconstruction (20 iterations) of a zero-energy point source with 20 cm standoff with (left) and without (right) a far-field approximation. White cross-hairs indicate the true source direction.

Figure 4.16 demonstrates how this effect weakens with increasing source distance by comparing the the reconstructed images with and without the far-field approximation through the normalized l^2-norm of the difference

$$\kappa_d = \frac{||\boldsymbol{\lambda}_d - \boldsymbol{\lambda}_d^{\infty}||_2}{||\boldsymbol{\lambda}_d^{\infty}||_2}, \tag{4.1}$$

where λ_d^∞ and λ_d are the reconstructed images at distance d with and without the far-field approximation, respectively. The difference quickly decreases with increasing source distance until roughly $60-100$ cm, where the far-field approximation becomes appropriate.

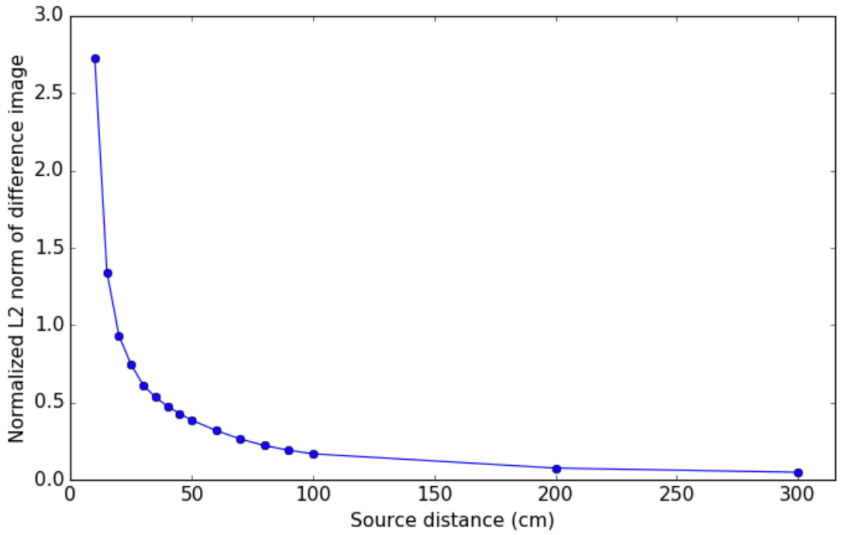

Figure 4.16. Normalized l^2-norm of the difference image with and without a far-field approximation with increasing source distance.

4.3.2 Static 3D Imaging

The 4π system response was simulated at 20 distances from 10 cm to 3 m and then interpolated to form a static 3D system response matrix. Depending on the size and discretization of the image space, this matrix can be too large to fit in the local graphics memory. Currently this forces the reconstruction to the CPU, resulting in slow reconstructions. This effect only becomes worse in the case of free-moving measurements. In an attempt to include some near-field effects with a more manageable system matrix, a far-field finite-detector approximation is used. Similarly to the approach used in previous sections, the far-field directional system matrix is passed to the GPU and the 3D system matrix is computed on the fly. However, the detector-to-voxel direction is computed from the crystal position in which the event occurred instead of the center of the sphere (i.e., assuming a point detector). This requires the crystal positions to be passed to the reconstruction and transformed into the global coordinate space for each event. This is straightforward in the static case as the global image space is defined in the local detector space. In the free-moving case, the coordinate transformation is done

with the translation and rotation matrices from the pose estimates. Note this approach will capture the finite detector effects in the near-field but will ignore magnification.

An experiment was performed to demonstrate the static near-field imaging capability of PRISM and the effects of using various near-field approximations in the reconstruction. A 20 μCi ^{241}Am source was placed 30 cm from the center of the PRISM sphere and data was collected for 5 min. The 3D image space was defined around the center of the detector, spanning 75 cm along each axis. The voxel size was set at 1.5 cm, with a total of $\sim 8 \times 10^5$ voxels. No SDF constraint was applied. At the time of the measurement, 83 detectors were operating sufficiently well at low-energies for the data to be included in the reconstruction. A 20 keV ROI was used around the 60 keV photopeak. Figure 4.17 shows the ML-EM reconstructions (75 iterations) using the far-field, far-field with finite-detector approximation, and interpolated near-field responses. The far-field response is unable to localize the source in 3D, with a hotspot localization error $>$ 45 cm. Both the finite-detector approximation and the interpolated near-field responses accurately localized the source (hotspot within one voxel).

To gain an understanding of the depth resolution with distance, this measurement was performed for a number of source standoffs. The 1D image slices along the source distance axis (Z) were then fit with exponentially modified Gaussian functions (due to the asymmetrical tailing on the far-side of the source) to extract a width parameter (σ_z). The measured width is plotted as a function of the source distance squared over the diameter of the detector (d^2/D) as this is a common metric used for depth camera characterization [126], [127]. This is done for both the far-field finite detection (FF-FD) approximation and the near-field (NF) images. The results are shown in Fig. 4.18. Depth resolution of FF-FD is generally observed to be worse than NF, as expected, but the degraded depth resolution may be acceptable given the CPU-limited near-field response reconstruction.

4.3.3 Free-moving 3D Imaging

The hand-held free-moving design of PRISM can overcome the inverse square law and get closer to sources than traditional static imagers. This facilitates efficient detection of weak sources as well as increased spatial resolution of 3D images. A source search scenario is considered here to demonstrate the near-field effects on free-moving 3D imaging as well as the computational burden of including varying degrees of the near-field response in the reconstruction. A 20 μCi ^{241}Am source was placed in a box on a bench in a small indoor laboratory space (see Fig. 4.19). PRISM-v0 (74 operational detectors) was walked up close to the suspected box and collected data from multiple positions and orientations to enhance the coded sampling. The total measurement time was 1.5 min (575 poses). The image space was discretized with 5 cm voxels (2×10^6 total voxels, 5×10^4 occupied voxels). A 15 keV ROI was used around the 60 keV photopeak energy

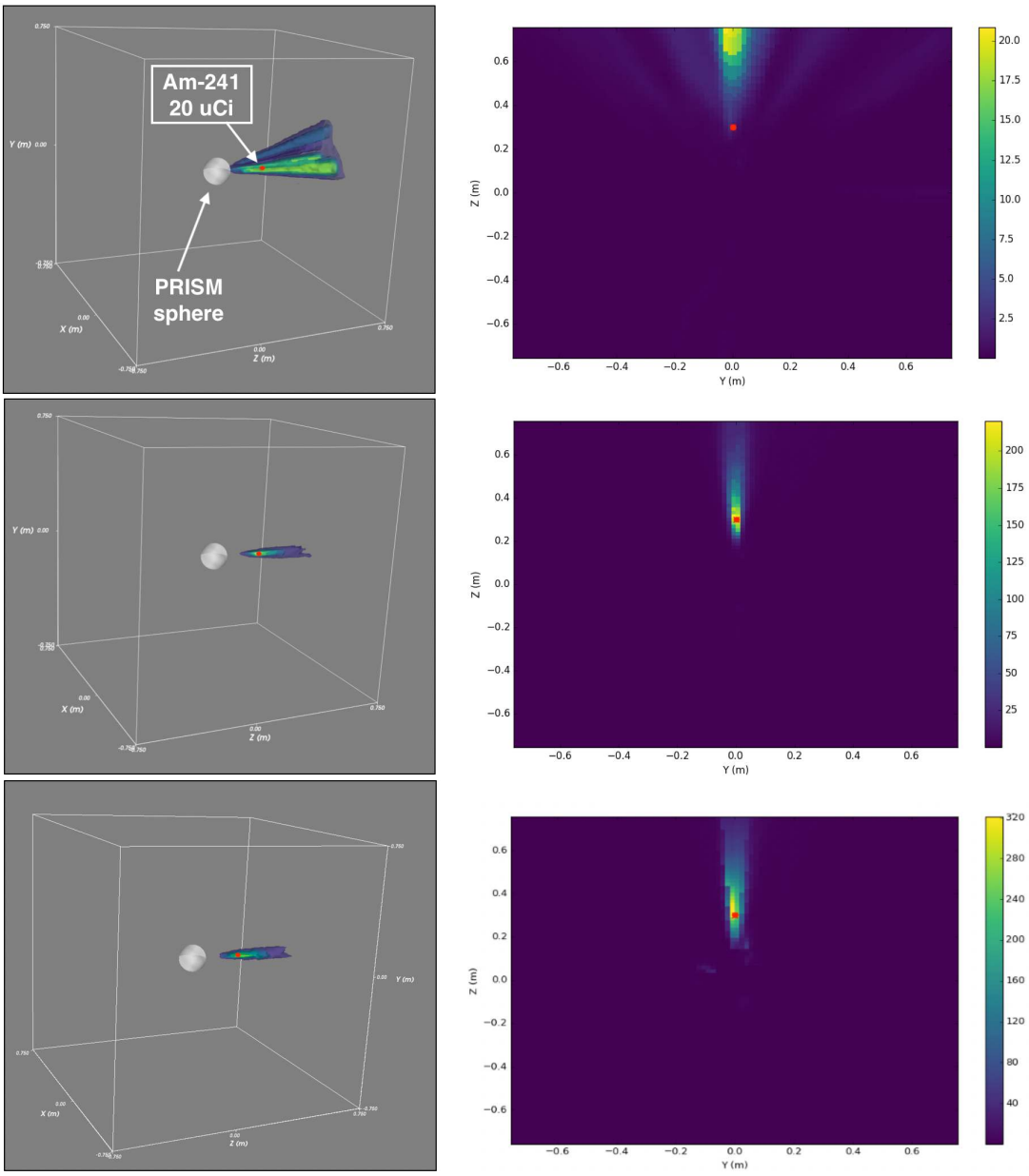

Figure 4.17. ML-EM reconstructions (75 iterations) using the far-field (top), far-field with finite-detector approximation (middle), and interpolated near-field (bottom) responses of a 20 μCi ^{241}Am source (marked with a red dot) placed 30 cm from the center of the PRISM sphere (shown to scale with a white sphere). The left column shows the reconstructions as 3D contour plots and the right columns shows 2D image slices through the X-axis. Measurement time was 5 min and voxel size is 1.5 cm.

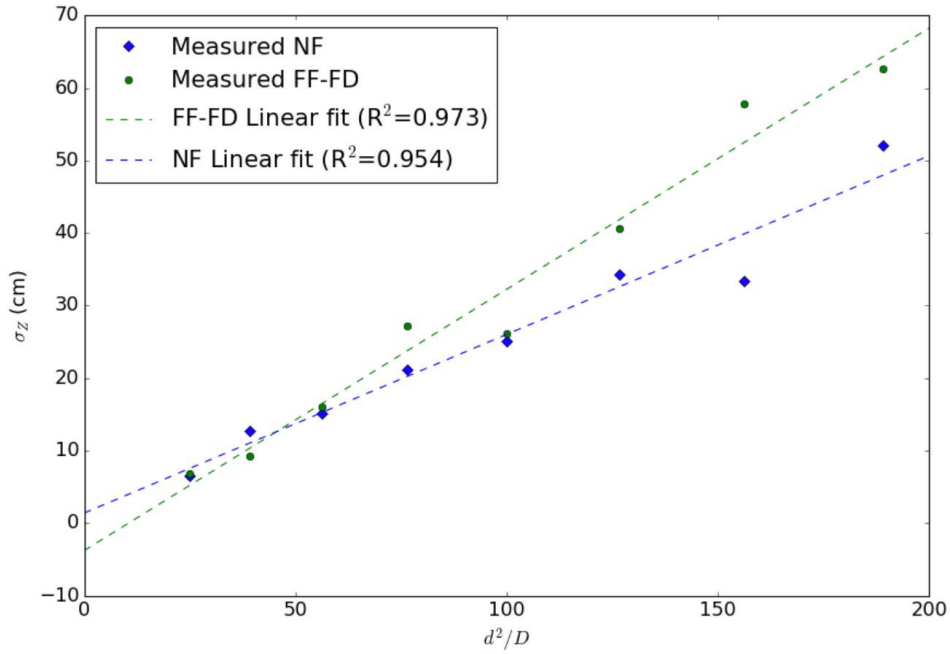

Figure 4.18. Depth resolution (FWHM of 1D depth image) as a function of source distance squared over the diameter of the detector for both the far-field with finite detector (FF-FD) approximation and the near-field (NF) reconstructions.

(11810 events). Twenty iterations of ML-EM were run with both the FF-FD and NF responses and the results are shown in Fig. 4.20. The FF-FD approximation localizes the source in the box to within several cm, but the reconstruction is noisy with image intensity on the back wall and orange storage bin on the lab bench. The hotspot is also slightly biased towards the corner of the box. The near-field response successfully removes the surrounding noise and more accurately localized the source. The CPU limitation of the near-field response, however, results in a $\mathcal{O}$(min) reconstruction time ($\sim$35$\times$ slower than FF-FD).

Additional work is required to progress towards a real-time near-field reconstruction approach. One approach would be to reduce the number of image voxels (i.e., constrain the near-field reconstruction only to a specific volume of interest). This space could be defined manually by the user or by thresholding a coarse image reconstruction using the far-field response. Another approach is a non-uniform discretization of the image space to force smaller voxels in the more sensitive regions of the image space and larger voxels in less sensitive regions. This focuses most of the voxels closer to the system where near-field effects can dominate and where spatial resolution is expected to be better. Additionally, compute cycles are not wasted on far voxels that are not expected to be resolved, given the angular resolution of the system.

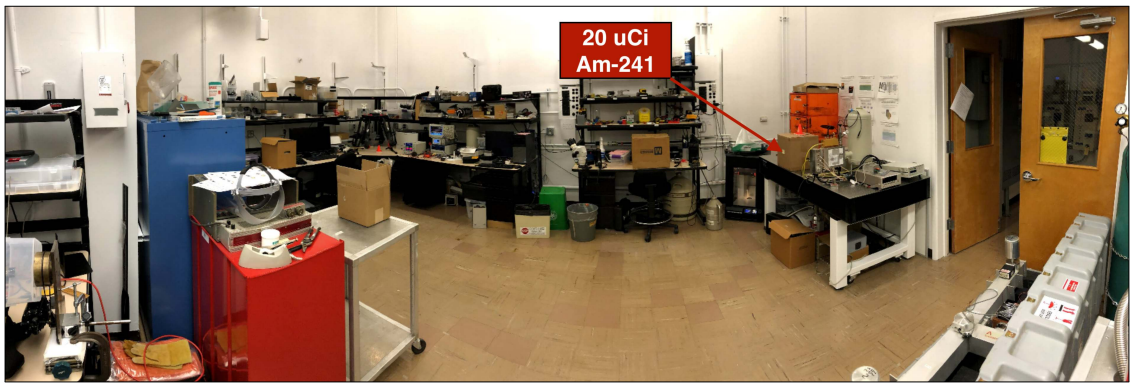

Figure 4.19. Free-moving near-field source localization scenario setup. A 20μCi ^{241}Am source is placed in a box on a bench in a small indoor lab space (marked with an arrow).

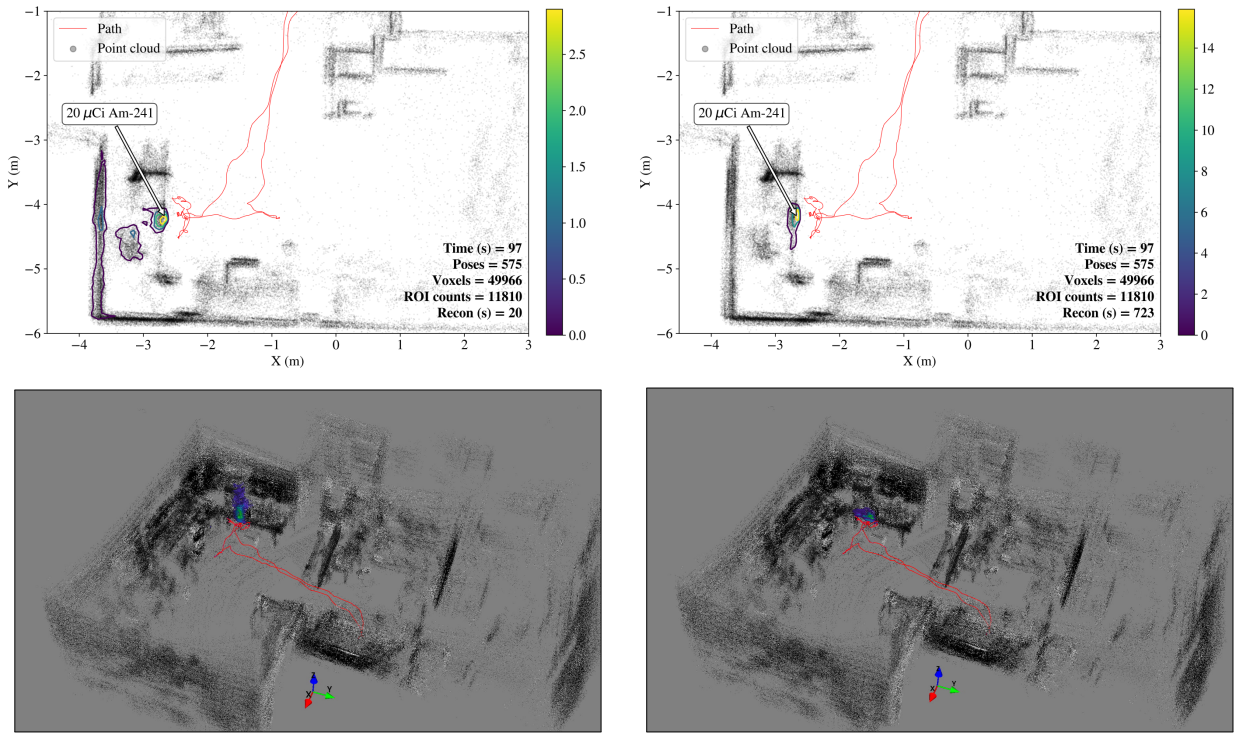

Figure 4.20. ML-EM reconstructions (20 iterations) of a near-field measurement using the far-field finite-detector (left) and near-field (right) system responses, visualized with top-down projections (top) and 3D contours (bottom).

4.4 Depth-of-Interaction Improvements

The read-out scheme of the CPG-CZT crystals in PRISM-v1 facilitates the event-by-event position estimation along the anode-cathode axis, or DOI. The DOI effectively increases

the information density in each interaction, providing more unique and high-frequency coding in the detection pattern across the sphere. In principle, the improved coding should improve the angular (and spatial in the free-moving 3D case) resolution of the system.

4.4.1 2D Imaging

Consider the measurement shown in Sec. 3.3.1 (20 μCi ^{133}Ba at 1 m standoff in front of PRISM-v1 on-axis). The zero-energy response was used to reconstruct a 2D image (20 iterations ML-EM) using a 20 keV ROI around the 81 keV photopeak. Figure 4.21 shows the reconstructed image in 4π (interpolated Cartesian projection) using simulated and measured data (20 min dwell), with the true source location marked with white cross-hairs. Separate images with and without DOI modeled in the response are also shown.

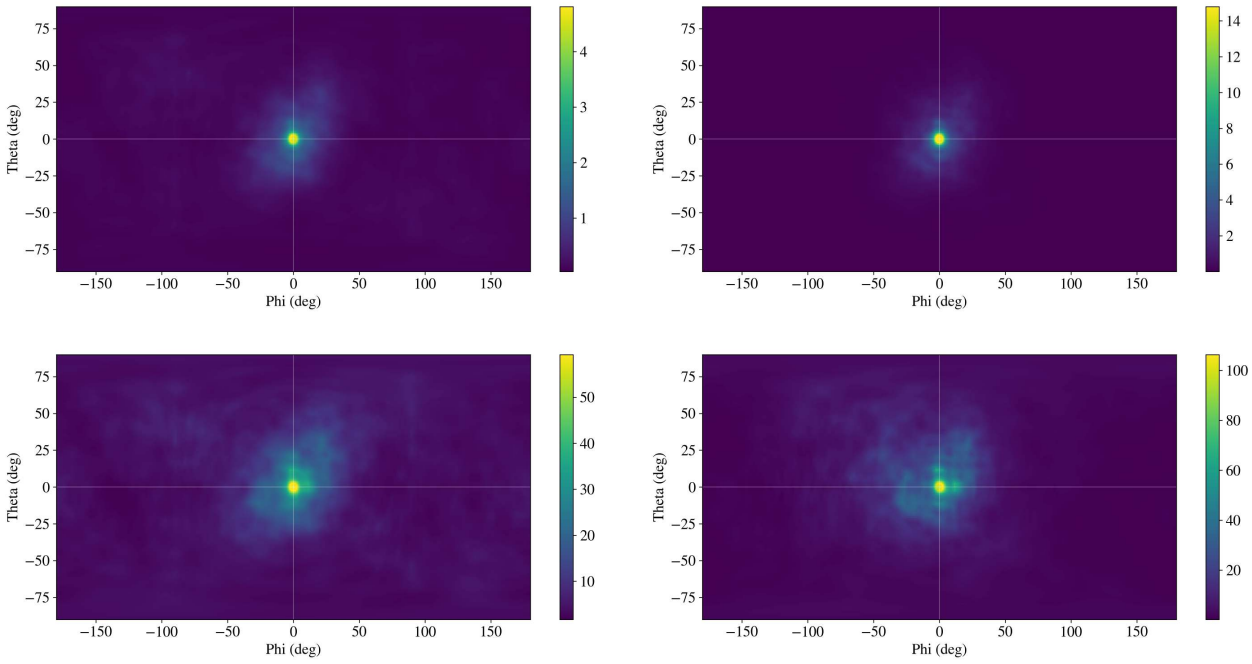

Figure 4.21. ML-EM reconstructions (20 iterations) of a simulation (top) and 20 min measurement (bottom) of a 20 μCi ^{133}Ba source on-axis with 1 m standoff. Reconstruction with (right) and without (left) the DOI response.

Image degradation is observed with the measured data, as expected, due to the presence of additional statistical and systematic noise in the data not modeled in the response (environmental radiation and scattering, down-scattering, electronic noise, etc.). Though in both cases, the DOI response results in an overall improved image quality.

One-dimensional image slices along the white cross-hairs ($\theta = 0$ and $\phi = 0$) are shown in Fig. 4.22, fit with Gaussian distributions to extract the FWHM. In both the simulation and measurement, the DOI response produces more peaked images with smaller FWHM.

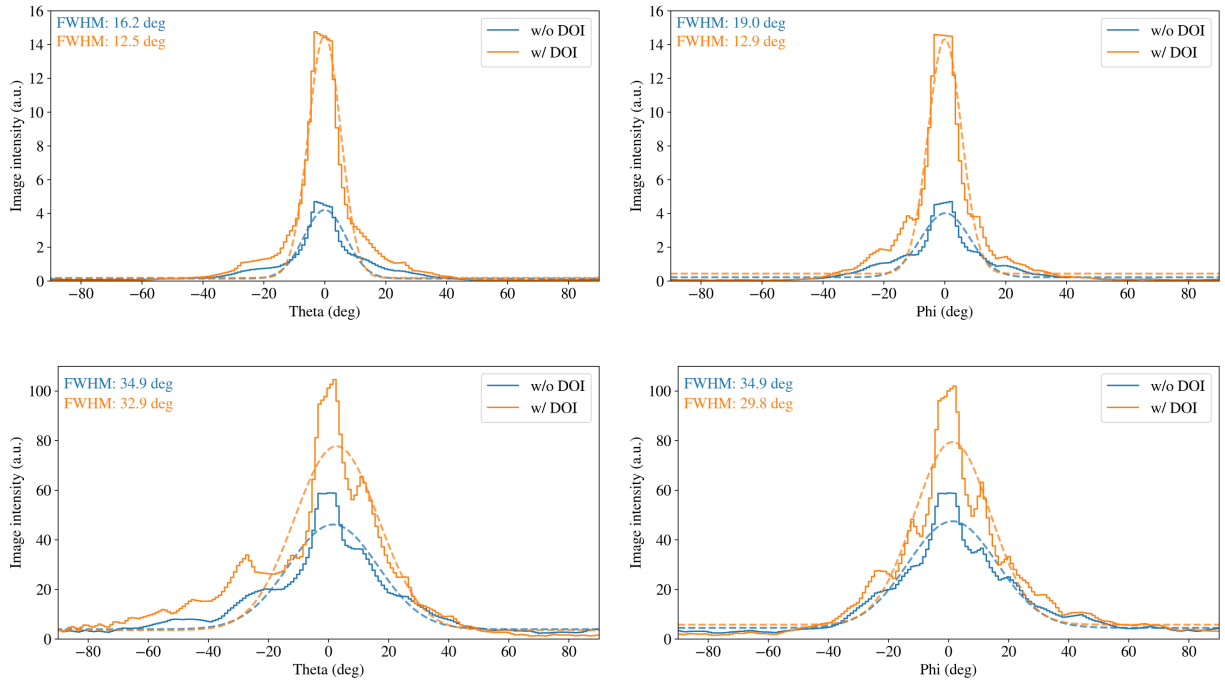

Figure 4.22. One-dimensional image slices along $\phi = 0$ (left) and $\theta = 0$ (right) in Fig. 4.21 - (top) simulation, (bottom) measurement. The images are fit with Gaussian distributions to extract the FWHM.

4.4.2 3D Imaging

The improvement to 3D imaging with DOI is first explored with a simulation of a short semi-circular path (134 poses) near a far-field zero-energy point-source (3.5 m standoff). Background and counting statistics were ignored. The orientation of the system was held fixed facing the source during the trajectory because the current configuration of DOI capable detectors in PRISM is such that imaging should be improved for directions in front of the system. The image space was uniformly voxelixed with 17 cm voxels around the path ($\pm$4 m in X and Y, $\pm$2 m in Z). Figure 4.23 shows the 3D image reconstructions (ML-EM 50 iterations) and the 2D image slices (YZ) along the peak intensity voxels, with and without DOI. The DOI reconstruction was successful in significantly increasing the localization precision ($\sim$15% decrease in spatial FWHM).

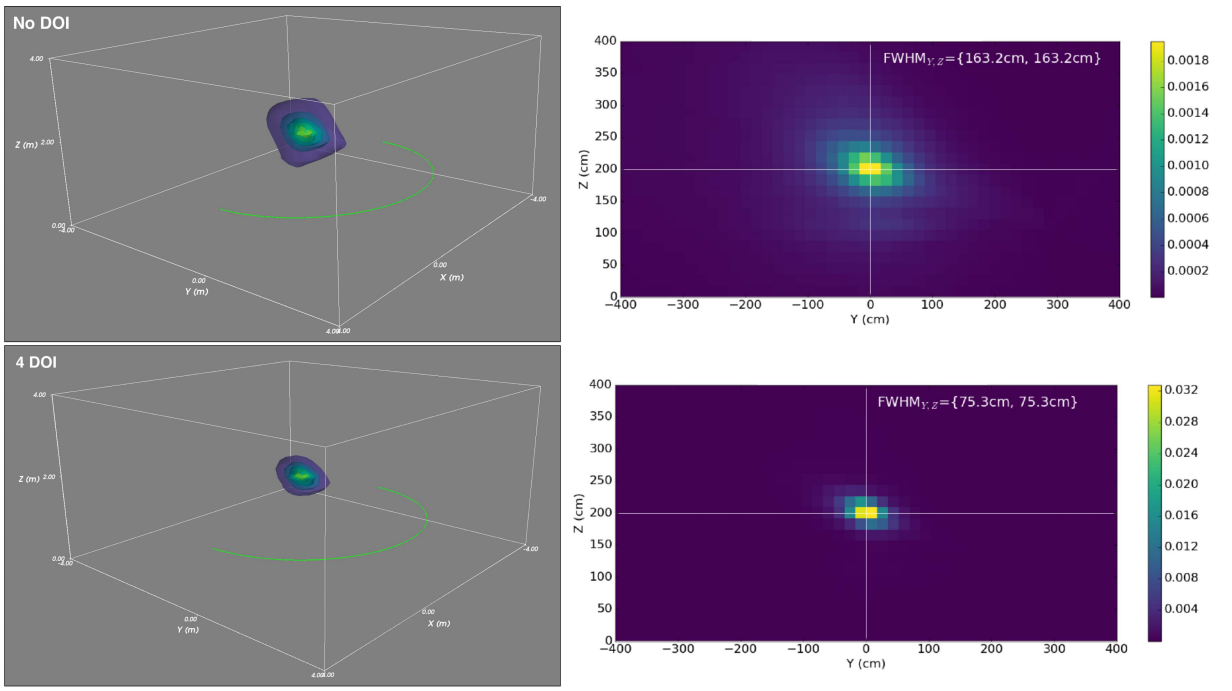

Figure 4.23. ML-EM reconstructions (50 iterations) of a simulated zero-energy point-source located 3.5 m from a semi-circle path (134 poses) shown in green without DOI (top) and with DOI (bottom). Two-dimensional Y-Z image slices along the peak intensity voxel are shown for each reconstruction.

Next, the PRISM-v1 system was slowly walked around a 20 μCi ^{133}Ba source in a stack of boxes in a small indoor laboratory space (in a similar semi-circle-type fashion as the simulations above). The entire measurement was approximately 1 min long (621 poses). The image space was voxelized with 10 cm voxels (6.8×10^4 occupied voxels). A 20 keV ROI was used around the 81 keV photopeak (9800 events). Figure 4.24 shows the top-down projection images (ML-EM 20 iterations) with and without DOI. No significant improvement to spatial resolution is observed, though the surrounding noise is reduced and, due to the count-conservation in ML-EM, more of the gamma-ray intensity is localized near the source. Note that only 39 detectors with cathode readouts were operational during this measurement. Future work includes the fabrication and integration of more DOI capable detector modules into PRISM-v1. This will increase the DOI sensitivity and the coded FOV. Further experimental characterization of the DOI response is also necessary to improve the system matrix currently used for reconstruction. This includes an understanding of how the depth resolution varies with depth and energy, edge and corner effects, and the impact of any dead layer associated with the coplanar grid electrodes [128], [129]. Finally, the inner (nearest cathode) depth bins contain the highest degree of coding because, at low-energies, the remainder

of the crystal acts as a significant attenuator and suppresses the broad response across the hemispherical section of 4π where the crystal is un-occluded by surrounding detectors. Therefore, low-energy events in the inner bin will most likely have originated from directions across the opposite side of the sphere, where the useful occlusion coding is present. These events could be more heavily weighted over other depth bins to further improve angular (and spatial) resolution.

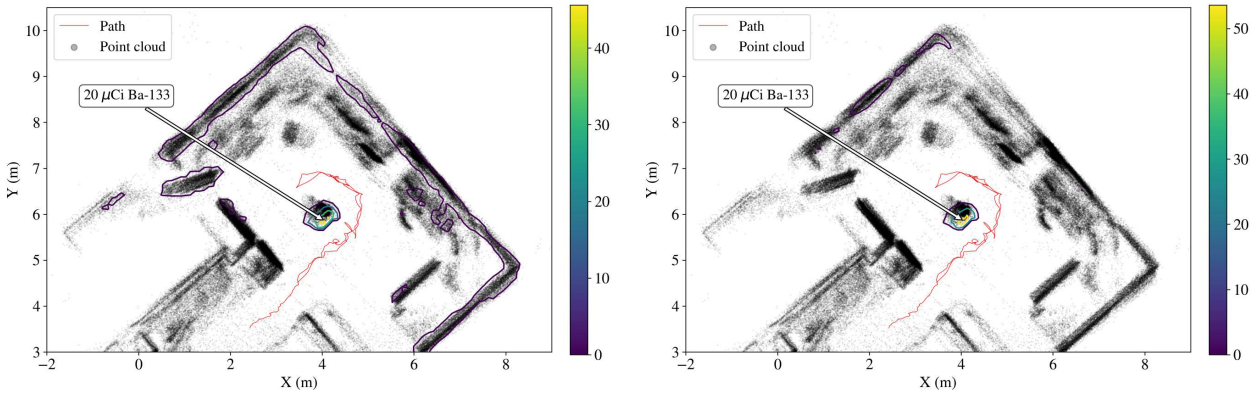

Figure 4.24. Top-down projection images (ML-EM 20 iterations) without (left) and with (right) the DOI response from a free-moving measurement around a 20 μCi ^{133}Ba source in a stack of boxes in a small indoor laboratory space.

4.5 Conclusions and Future Work

The use of free-moving systems and the integration of contextual sensors and SDF have been shown to improve source localization and distribution mapping capabilities for applications ranging from nuclear security to consequence management. This work is the first to experimentally demonstrate an approach to low-energy real-time 3D gamma-ray imaging with SDF and GPU-accelerated list-mode ML-EM using a hand-held CZT-based dual-mode omnidirectional active spherical coded mask system. Source localization of unshielded point-sources was successful over energies from 60−356 keV, activities from 20−5000 μCi, and areas of size 70−4500 m^2. Small indoor and wide-area outdoor scenarios were considered and measurements were all $<$ 3 min, exemplifying the capability of free-moving systems to overcome the inverse square law for efficient localization. Additional work is required to demonstrate the ability to image shielded or extended sources and map distributed sources in real-time.

In its current form, the image reconstruction slows with the addition of more contextual and gamma-ray data. Progress towards a scalable image reconstruction approach

will be continued in future work. This includes the implementation of a real-time probabilistic tri-state occupancy model and the removal of transient points in the point cloud. The uniform spatial voxelization scheme may not be optimal and other approaches will be explored, including adaptive non-uniform spatial discretization based on distance of approach and overall sensitivity (distant or low sensitivity voxels can be made larger, while closer voxels made smaller for increased spatial resolution).

Near-field effects were modeled with a perspective projection in the zero-energy OpenGL simulation. PRISM successfully demonstrated artifact-free 3D source localization when the source was within 30 cm of the detector enclosure, in both the static and free-moving cases. The size of the full near-field system response limited the reconstruction to the CPU, resulting in non-real-time reconstruction speeds. A far-field finite-detector approximation was introduced to include some near-field effects and maintain the real-time reconstruction performance.

The design of the CPG-CZT modules and the PRISM-v1 system facilitates event-by-event DOI estimation. The increase in information density of each event and the high-frequency features in the low-energy response were shown to improve the quality of the active coded mask imaging modality in both 2D and 3D. In addition to integrating more DOI-capable detector modules into the system, a thorough experimental characterization of the DOI response will be necessary to continue to improve image quality.

Chapter 5

APSL

Gamma-ray imaging attempts to reconstruct the spatial and intensity distribution of gamma-emitting radionuclides from a set of measurements. Generally, this problem is solved by discretizing the spatial dimensions and employing the ML-EM algorithm, with or without some form of regularization. While the generality of this formulation enables use in a wide variety of scenarios, it is susceptible to overfitting, limited by the discretization of spatial coordinates, and can be computationally expensive. In this chapter, a novel approach to 3D gamma-ray image reconstruction is presented for scenarios where sparsity may be assumed, for example radiological source search. First a Point-Source Localization (PSL) approach is formulated as an optimization problem where both position and source intensity are continuous variables. This formulation is extended and generalized to an iterative algorithm called Additive Point-Source Localization (APSL) for sparse parametric 3D image reconstruction. A set of simulated and experimental source search scenarios using a single non-directional detector are considered, finding improved image accuracy and computational efficiency with APSL over traditional grid-based approaches.

It is important to make clear the distinction of the approach described in this chapter and the one in Ch. 4. The PRISM system represents a state-of-the-art free-moving gamma-ray imaging system that can be used for source localization and distribution imaging across a broad range of energies using the active coded mask and Compton imaging modalities. Though it is still solving an underdetermined inverse problem, the superior imaging capability ($< 10^\circ$ angular resolution) and SDF can significantly constrain the reconstruction and produce high-quality 3D images and accurately localize compact sources in real-time with ML-EM, as demonstrated in Ch. 4. This chapter addresses the radiological source search scenario from a more general perspective, beginning with a free-moving non-directional single detector system using only the inverse square law for source localization. The results of this work show that, under certain assumptions, this new approach can outperform ML-EM and can produce accurate 3D

multi-point-source localization estimates without the use of sophisticated and expensive imaging systems.

5.1 Introduction

The reconstruction of the spatial and intensity distribution of gamma-emitting radionuclides from a set of Poisson-distributed measurements is employed for applications ranging from medical imaging to nuclear security. In the case of medical imaging, measurements are often from many static detectors viewing a stationary volume, whereas in the case of nuclear security, measurements may be from one or several detectors that have moved through an environment. In both cases, the position and orientation (pose) of the detectors relative to the scene must be known; in the former the detector geometry is known *a priori*, however in the latter case pose information may be provided by geospatial positioning (GPS), an inertial navigation system (INS), or SLAM [34], [35].

Here the subset of scenarios where the source-term may be assumed sparse is considered. In the case of radiological source search, the task is often to localize and quantify the activity of one or more compact sources of gamma-ray radiation. Under a point-source assumption, previous work in static regularly-spaced 2D detector networks has considered techniques such as triangulation using least-squares [130], hybrid grid-based maximum likelihood estimation and expectation maximization [131], and regularized ML-EM pre-conditioned Fisher's scoring iterations [132]. In the space of adaptive networks or moving detectors, approaches such as two-stage algorithms [133], adaptive likelihoods [134], and sequential Bayesian estimation using particle filters [135] have been explored. Other approaches in the literature include grid-refinement and iterative pruning [136]–[139].

The existing approaches tend to localize a known number of sources (primarily in 2D), assume background is known or can be estimated prior to source localization, and rely on the discretization of the spatial and intensity domains either directly or for initialization of a direct solver. Grid-based methods are limited in accuracy and can be computationally intractable when searching for multiple sources [132]. Moreover, methods developed in 2D may not be easily extensible to 3D localization, due to increasing degeneracy and non-convexity in the solution space.

More recent work has approached the problem in the maximally general case (unknown background and no priors on the source distribution), using ML-EM to perform 2D and 3D image reconstruction of sparse and distributed sources with a variety of static and free-moving detector systems including hand-held [19], [30], [33], ground-vehicle [70], [140], and airborne platforms [29], [141]. Various regularization approaches to impose assumptions about the source distribution have also been studied. These include sparsity priors such as those in [6], [92], [142] and scene data constraints such as those

presented in Ch. 4 and [19], [36], [43]. While the generality of the ML-EM formulations enable use in a wide variety of scenarios, it is also susceptible to overfitting, limited by spatial discretization, and can be computationally expensive and memory intensive.

First, the general ML-EM approach is considered in a fully discretized (not SDF-constrained) 3D space subject to Poisson gamma-ray counting statistics, with and without sparsity regularization, and the limitations of such an approach is demonstrated. A reformulation of the problem is then proposed to one in which the source model is confined to a single voxel, i.e., PSL. This approach is similar to prior work [131], though here reconstruction is in 3D and background is included as a free parameter. The PSL problem in which both the spatial and intensity domains are continuous is then considered. Finally, the continuous PSL formalism is extended for general sparse image reconstruction, or APSL, where the image is the sum of multiple point-sources whose position and intensity are continuous in nature. APSL mitigates overfitting in its iterative bottom-up nature and statistically-founded stopping criteria and, because of the inherent point-source assumption and continuous variables, results in images with improved accuracy and interpretability as compared with traditional grid-based approaches. Furthermore, APSL offers this enhanced performance at a reduced computational burden.

This chapter serves as an introduction and simple demonstration of the APSL approach. The formalism is general in nature in that it can search for an unknown number of sources in an unknown background environment using static and dynamic detectors or detector arrays. To demonstrate the concept here, a simulated source search scenario comprising a single free-moving detector with uniform directional sensitivity whose poses are derived using SLAM is considered. The detector poses are subject to arbitrary rotations and translations in 3D as if being carried by a human operator. An isotropic detector was used for simplicity and clarity, though this case is challenging for both localization and quantification due to the inability to break solution degeneracy. Directional detectors and detector arrays can aid in breaking degeneracy, but complex angular responses may present additional challenges for the non-convex Poisson likelihood optimization approaches developed here and are subject to future work. The approach is then experimentally demonstrated with a source search scenario using a nearly isotropic free-moving $LaBr_3$ detection system with a contextual sensor package including LiDAR and IMU for SLAM and SDF.

The structure of this chapter is as follows: the traditional and sparsity regularized ML-EM approaches for a simulated source search scenario are outlined in Sec. 5.2−5.3, followed by the discrete and continuous PSL formalisms on the same scenario in Sec. 5.4. The APSL algorithm is demonstrated on a simulated multi-source localization scenario in Sec. 5.5 and experimental APSL results are shown in Sec. 5.5.2. A summary and future work are presented in Sec. 5.6.

5.2 Maximum Likelihood Expectation Maximization

To demonstrate the limitations of ML-EM for sparse image reconstruction, a simulated isotropic (non-directional) 100% efficient detector (effective area of 5 cm^2) is simulated at points along a short experimental human-walked trajectory (270 poses) in a 20 m^2 area. The trajectory was tracked with a LiDAR-IMU sensor system and the Google Cartographer SLAM algorithm [74]. The elevation (Z) of the path varied slightly, roughly $\pm$30 cm about the XY plane (Z = 0). The choice of a simulated detector and measured trajectory was made here for the ease of single (and multiple) point-source injection studies while capturing the variation in detector-source distance (in 3D) observed in real free-moving source search scenarios. A 5 μCi point-source was simulated in the XY plane with a closest approach of $\sim$45 cm and without regard to any spatial discretization for ML-EM. A background count rate of 100 counts per second was assumed and each simulated measurement had a duration of 0.1 s.

The top pane of Fig. 5.1 shows the path of the detector in the XY plane with heading described with arrows, colorized with the total counts at each measurement. Larger arrows indicate a faster speed in the trajectory, though the simulations were done statically at each position. The source position is shown with a red "x". The middle pane shows the position of the detector in Z at each measurement. Notice the detector was positioned at Z ≈ -14 cm for the majority of the path near the source. The bottom pane shows the simulated signal and background ("bkg") at each measurement.

A non-directional detector was used here to limit the reconstruction to proximity imaging (relying only on flux modulation from the inverse square distance to the source) and focus on the development and performance of the algorithms that follow.

The 3D image space was discretized into cubic voxels (10 cm), with a Z extent from -3 to 3 m (total of $\sim 1.28 \times 10^5$ voxels). The image space was centered around Z = 0 as the path of the detector was roughly centered around this plane. The point cloud was not used to constrain the image space. This will be considered in Sec. 5.5.2.

Figure 5.2 shows a slice of the ML-EM reconstruction (20 iterations), in $\log_{10}$ scale, along the XY plane (Z = 0). Image estimates are represented by $\hat{\mathbf{w}}$ here to differentiate from the results in Ch. 4 (which used $\hat{\boldsymbol{\lambda}}$). The number of iterations reflects a 1.5×10^{-3} fractional change in negative log-likelihood and thus a reasonable estimate for convergence. Twenty iterations were used in all subsequent reconstructions for consistency (in computation) between methods. Also shown is the forward projection of the ML estimates ($\hat{\bar{\boldsymbol{x}}} = \mathbf{V} \cdot \hat{\mathbf{w}} + \hat{b}\boldsymbol{t}$) compared to the measurement, $\boldsymbol{x}$. The optimized negative log-likelihood (hereby referred to as the "loss") of the fit is shown in the upper right corner of the lower plot. As the problem is underdetermined, the ML-EM solution significantly overfits, placing intensities in many voxels across the whole image space, ultimately fitting to the noise. This effect increases with additional iterations.

The system matrix calculation and ML-EM reconstruction were run on a quad-core

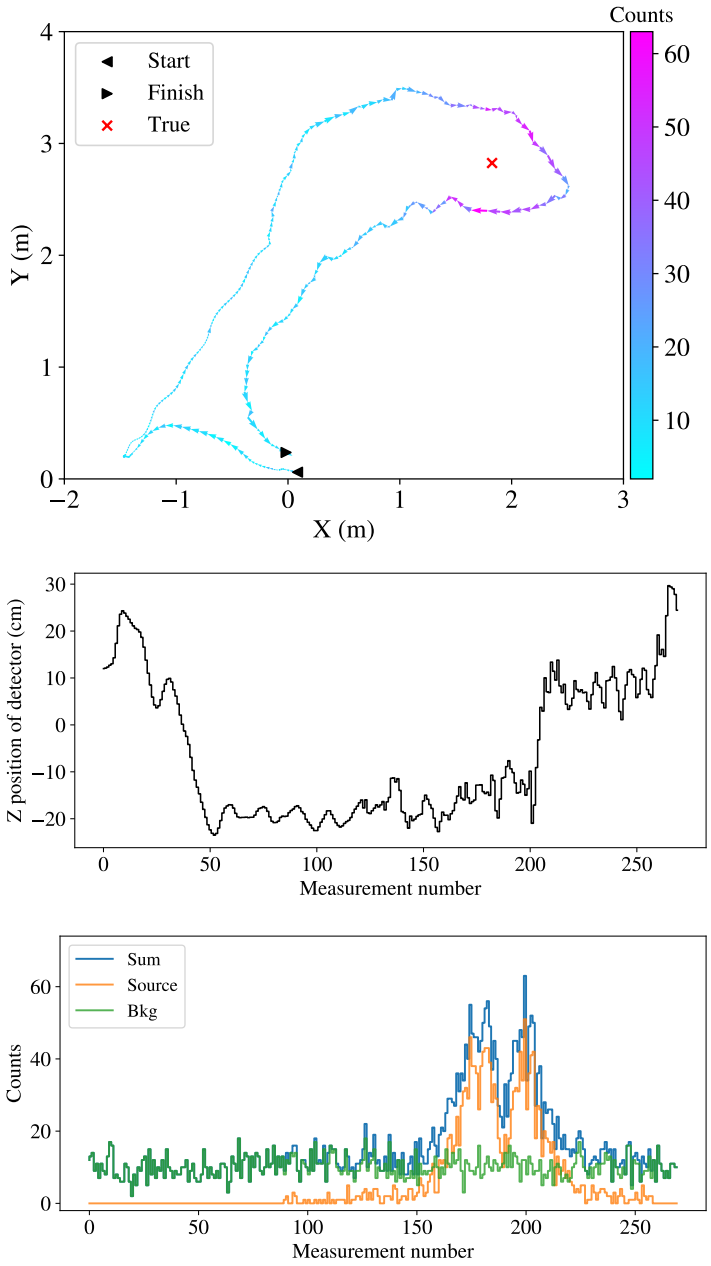

Figure 5.1. (Top) Measurement path in the XY plane, colorized with the total simulated counts at each measurement. The path is a human-walked trajectory tracked with a SLAM system. The simulated measurements remained static at each position for a duration of 0.1 s. The source position is shown with a red "x". (Middle) Measurement path along the Z dimension. (Bottom) Simulated signal and background at each measurement.

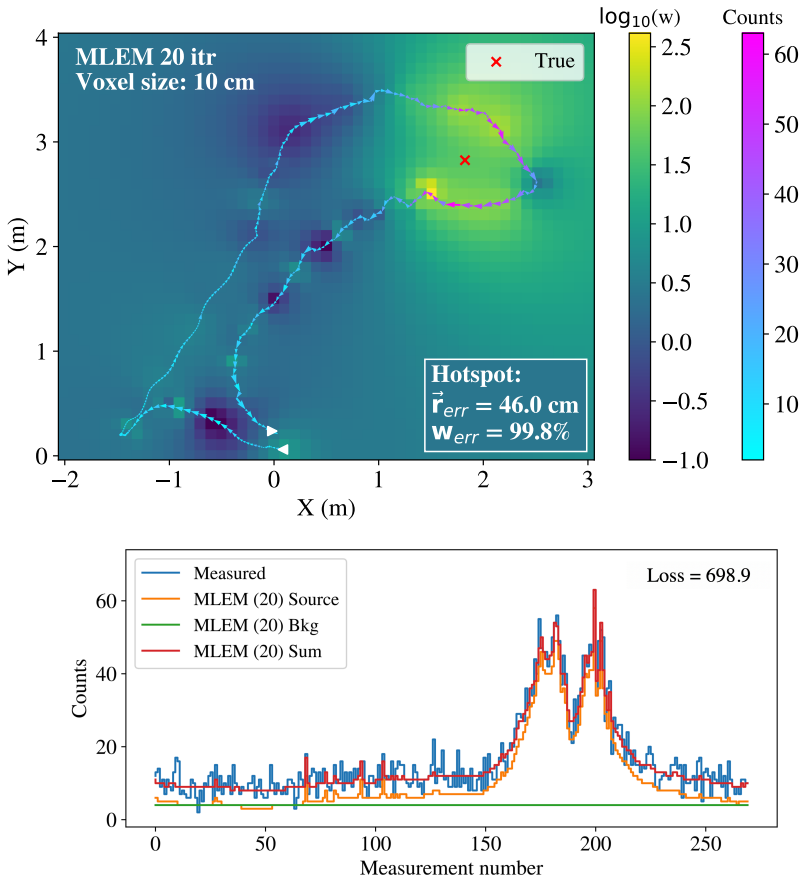

Figure 5.2. (Top) ML-EM reconstruction $\hat{\mathbf{w}}$ (20 iterations), in $\log_{10}$ scale, of a slice through the XY plane (Z = 0). A 5 μCi source was placed at the red "x" at Z = 0. (Bottom) Forward projection of ML estimates into count space $\hat{\bar{x}}$ compared to the measurement $\boldsymbol{x}$. The total negative log-likelihood of the fit is shown in the upper right corner.

2.7 GHz Intel Core i7 CPU with runtimes of 22 s and 0.7 s, respectively. The same hardware was used for all reconstructions in this chapter. ML-EM was performed on the CPU to fairly compare against the CPU-based PSL and APSL implementations. GPU acceleration techniques for APSL are left for future work.

While ML-EM successfully minimizes the loss, the solution fails to localize the source and places most of the intensity near the path of the detector. The bias towards the measurement path results in a localization error (Euclidian distance between the true location and the highest intensity voxel center, $\vec{r}_{\text{err}} = ||\vec{r}_{\text{true}} - \vec{r}_{j=\text{argmax}(\hat{\mathbf{w}})}||$) of 46 cm ($\sim$100% of closest approach). To compensate for a much closer source, the individual source intensities are much lower than the true value. However, the ML-EM background estimate is considerably lower than the true rate, resulting in more source intensity being placed throughout the 3D image space (total of $\sim$19 μCi).

5.3 Maximum A Posteriori

This section explores the use of sparsity priors to regularize ML-EM towards point-source solutions. Figure 5.3 shows the MAP solution $\hat{\mathbf{w}}$ (20 iterations of EM) to the data in Fig. 5.1 using the log prior defined in Eq. 2.33 ($\delta = 1$ and $\rho = 0.01$) as well as the comparison of $\hat{\bar{x}}$ to the measurement. The reconstruction runtime and likelihood convergence were similar to ML-EM.

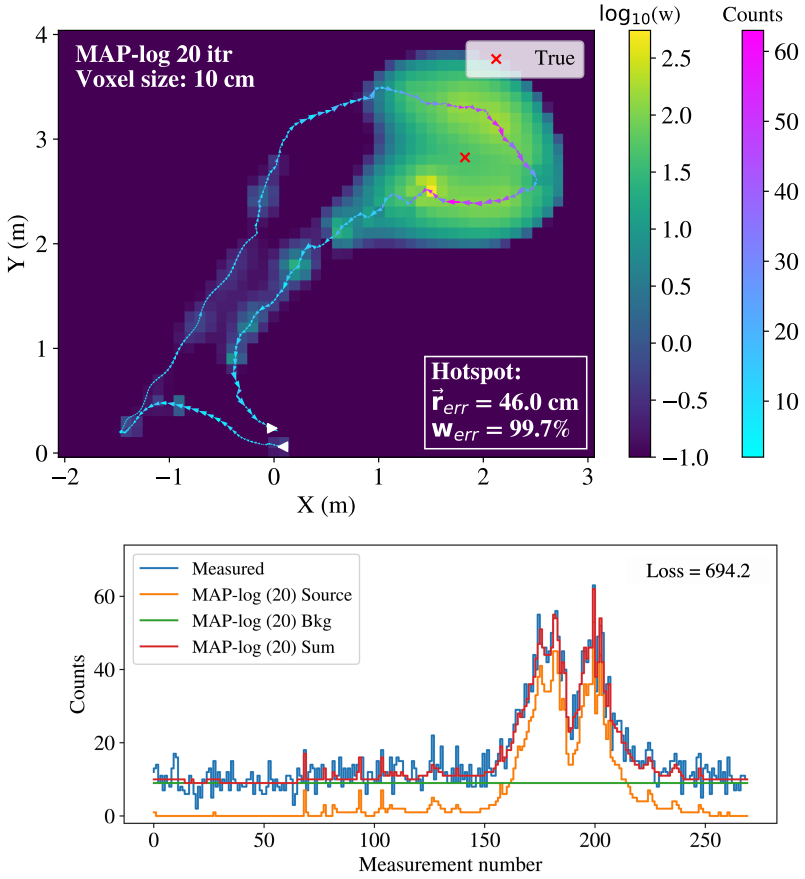

Figure 5.3. (Top) MAP reconstruction $\hat{\mathbf{w}}$ (20 iterations of EM) with the log prior in Eq. 2.33 ($\delta = 1$ and $\rho = 0.01$), in $\log_{10}$ scale, of a slice through the XY plane. (Bottom) Forward projected mean-rates $\hat{\bar{x}}$ compared to the measured signal x. The total negative log-likelihood of the fit is shown in the upper right corner.

Hyperparameter optimization was done with a coarse grid search over δ and ρ ($\delta, \rho \in 10^{\{-3,-2,-1,0,1,2\}}$), maximizing the fraction of the intensity near the source (within half the distance of closest approach). The optimal parameters were below the suggested range of δ and ρ in [92] due to the difference in scale in the data. While the solution is successful in suppressing much of the activity surrounding the trajectory and producing

a better background estimate than ML-EM, the log prior still suffers from overfitting, fails to localize the source, and biases the intensity near the path.

Figure 5.4 shows the MAP solution $\hat{\mathbf{w}}$ (20 iterations of EM) to the data in Fig. 5.1 using the gamma prior defined in Eq. 2.34 ($\alpha = 1.01, \beta = 1$, and $\rho = 10^{-3}$) and the fit in count space. Again, the reconstruction runtime and likelihood convergence were similar to ML-EM. A grid search approach similar to above was used for hyperparameter optimization (α, β, and ρ). Similar to the log regularizer, the gamma prior suppresses intensity surrounding the trajectory, though it still places intensity along the track and fails to localize the source.

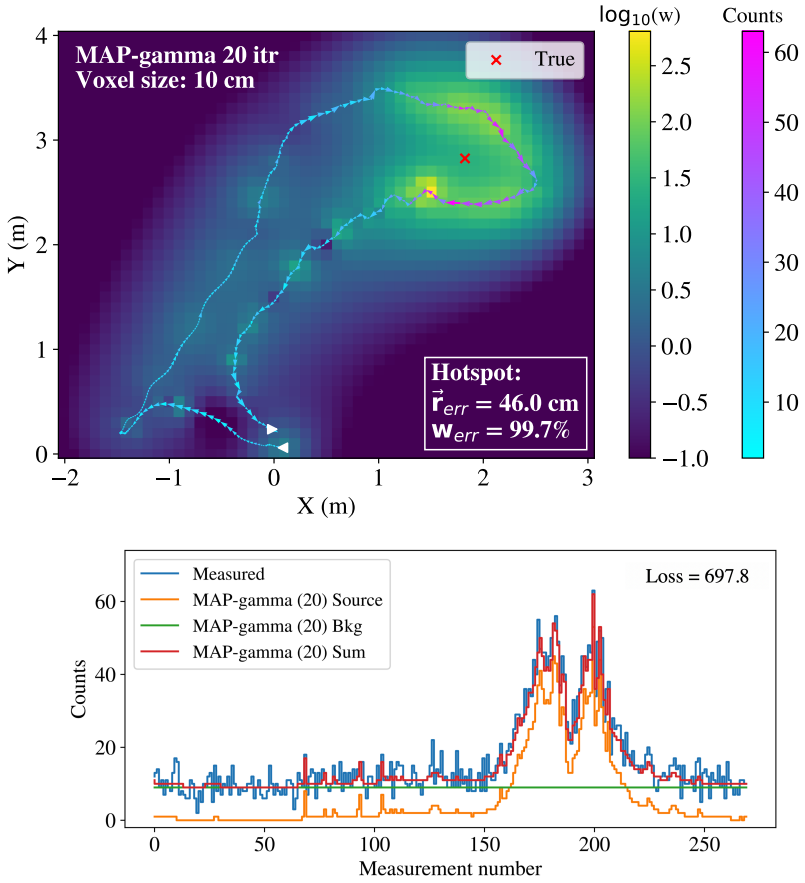

Figure 5.4. (Top) MAP reconstruction $\hat{\mathbf{w}}$ (20 iterations of EM) with the gamma prior in Eq. 2.34 ($\alpha = 1.01, \beta = 1$, and $\rho = 10^{-3}$), in $\log_{10}$ scale, of a slice through the XY plane. (Bottom) Forward projected mean-rates $\hat{\hat{x}}$ compared to the measurement. The total negative log-likelihood of the fit is shown in the upper right corner.

In both cases presented here, regularizing the reconstruction using prior information fails to improve source localization and quantification. Ultimately both problems are still severely ill-posed. However, if the unknown source is known to be a point-source

(i.e., a source occupying a single voxel), such as in radiological source search, significant improvements can be made by reformulating the optimization problem in Eq. 2.25.

5.4 Point Source Localization

5.4.1 Discrete Space

Under the point-source assumption, the optimal background rate and point-source intensities for each voxel can be solved using ML-EM with the following replacements

$$\begin{aligned}
\bar{\boldsymbol{x}} &\Rightarrow \overline{\boldsymbol{X}}^{[I\times J]} = [\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_J]\,, \\
\mathbf{w} &\Rightarrow \mathbf{W}^{[J\times J]} = \mathrm{diag}(\mathbf{w})\,, \\
\boldsymbol{x} &\Rightarrow \boldsymbol{X}^{[I\times J]} = [\boldsymbol{x}, \ldots, \boldsymbol{x}]\,, \\
b &\Rightarrow \boldsymbol{b}^{[1\times J]} = [b_1, \ldots, b_J]\,, \\
\boldsymbol{t} &\Rightarrow \boldsymbol{T}^{[I\times J]} = [\boldsymbol{t}, \ldots, \boldsymbol{t}]\,.
\end{aligned}$$

Equation 2.18 then becomes

$$\overline{\boldsymbol{X}} = \mathbf{V}\cdot\mathbf{W} + \boldsymbol{b}\odot\boldsymbol{T}. \tag{5.1}$$

The voxel in discrete space that alone best describes the data (minimum loss) may then be identified by solving Eq. 2.25 for each voxel. This can be thought of as individual ML-EM problems in each voxel, each solving for the optimal weight (and background) that best explains the measured data. While Eqs. 2.28 and 5.1 can be refactored for computational efficiency, the reconstruction will be slower than traditional ML-EM because more computations are required. In the point-source scenario, an analytical solution exists for the source intensities in the absence of background

$$\mathbf{W} = \boldsymbol{X}^T\cdot\mathbf{1}\,/\,\mathbf{V}^T\cdot\mathbf{1}\,. \tag{5.2}$$

This solution can be computed exactly with little computational power, however it does not hold in the case of the unknown background rate presented here. The analytical solution is used as an improved initial image (over a flat image) to increase the speed of convergence in the iterative approach. While only 5 iterations were required for the minimum loss voxel to achieve the same likelihood convergence as the ML approaches above, this was not true in every voxel in the image space. Therefore 20 iterations were used to be consistent with the other approaches.

Figure 5.5 shows the ML-EM loss for each single-voxel model of a slice through the XY plane. The reconstruction runtime was 25.2 s. The distribution agrees with intuition as the detector is non-directional, leading to many voxels outside of the track that can fit sections of the data (one of the two peaks in the data) albeit with a higher loss. It is only

the voxels near the true location that capture both peaks in the data, reflecting the lower loss.

The minimum loss voxel was one voxel above the true source location in Z (resulting in the position error $\sim$10 cm). This is not surprising as a non-directional detector was used and there was little change in elevation over the measurement path. Therefore some degree of degeneracy may exist in the maximum likelihood solution. Note the activity estimate for a point-source in this voxel was 4.7 μCi ($<$ 10% error).

While the minimum loss voxel can be used directly for the single-source localization (assuming the detection of the source above background has already been done with some other means, e.g., gross counts or spectral), spatial confidence intervals can be computed to bound the true source location with some degree of certainty (e.g., 95%). Under the assumption that the minimum loss voxel contains the true source location, confidence intervals can be computed around the voxel by comparing to the surrounding likelihoods with a likelihood ratio test.

The comparison of each loss ℓ_j to the minimum loss ℓ_{min} remains in the interior of the parameter space, and thus Wilks' Theorem [143] states that the test statistic of twice the negative log-likelihood difference

$$z = 2(\ell_j - \ell_{\text{min}}), \tag{5.3}$$

will be asymptotically distributed like χ^2_k, where the number of degrees of freedom k is the difference in the number of free parameters between ℓ_{min} and ℓ_j. The parameters of ℓ_{min} are all fixed, resulting in a difference of 5 free parameters (one source intensity, one background rate, and three spatial coordinates of the source), therefore $z \sim \chi^2_5$. For a system of N_d detectors, the solution would generalize to $k = 4 + N_d$.

Note that with degenerate maximum likelihood estimates (as can be the case with non-directional systems and limited movement in one or more of the spatial dimensions) the number of degrees of freedom $k \leq 4 + N_d$. Simulations or resampling methods would be required to determine the shape of the distribution and the true value of k (see Fig. 5.6), which may not be feasible in an operational search setting. However, assuming a larger k will always result in a more conservative confidence interval. While some degree of degeneracy may exist in this problem presented here, a value of $k = 5$ was used for the single detector system.

The equivalent Gaussian sigma (z-score) of a random variable x distributed like χ^2_k is

$$\text{z-score} = \sqrt{2}\ \text{erf}^{-1}\left[\Phi_{\chi^2_k}(x)\right] \tag{5.4}$$

where $\text{erf}^{-1}(\cdot)$ is the inverse error function and $\Phi_{\chi^2_k}(\cdot)$ is the cumulative distribution function (CDF) of the χ^2_k distribution. The middle image of Fig. 5.5 shows the confidence intervals around the minimum loss voxel in units of z-scores (zoomed in near the source

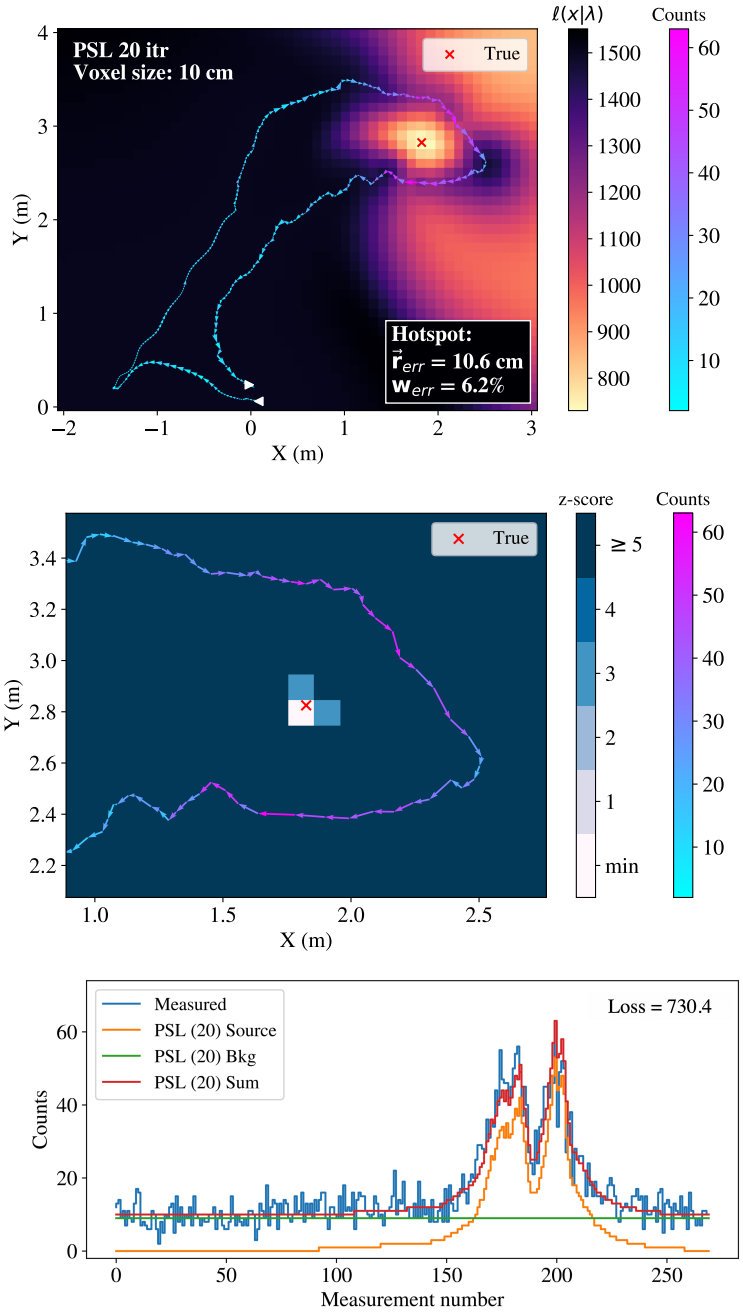

Figure 5.5. (Top) ML-EM loss (20 iterations), in $\log_{10}$ scale, for each single-voxel model (PSL) of a slice through the XY plane. (Middle) Likelihood ratio test statistic of the loss' to the minimum loss, represented with a z-score (zoomed in near the true source location). (Bottom) Forward projected mean-rates of the minimum loss single-voxel model $\hat{\hat{x}}$ compared to the measurement.

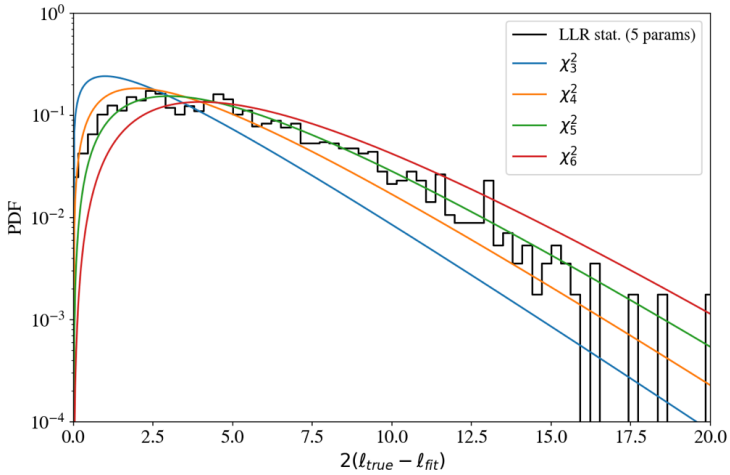

Figure 5.6. Distribution of the Wilks' test statistic (twice the negative log-likelihood difference) fit with χ^2_k for a simulated resampling of the problem posed in Fig. 5.5.

location). It is clear that the localization is sharply peaked with high confidence ($\geq 5\sigma$), demonstrating the superior localization capabilities of the approach.

The calculation of confidence intervals for the source intensity is more complex as it is coupled to the reconstructed position and background estimates. A more appropriate approach in this case would be to compute covariances from the Fisher information matrix (Cramer-Rao bound) or by determining the distributions empirically by gridding over intensity and background for the positions within some given confidence limit. However, this is outside the scope of this work.

The bottom image of Fig. 5.5 shows the minimum loss single-voxel model fit in count space. The solution is no longer overfit (resulting in a larger loss compared to the ML-EM and MAP approaches). The simple reformulation of the problem drastically improves both the localization and quantification performance, though, if the reconstruction is in discrete space, at the cost of a higher computational burden.

5.4.2 Continuous Space

The minimization of Eq. 2.16 can be reformulated as an optimization problem not only in continuous intensity, but also in the continuous 3D spatial coordinates of the point-source ($\vec{r}_s$)

$$\underset{(\mathrm{w}_s, \vec{r}_s, b)}{\operatorname{argmin}} \; \ell(\boldsymbol{x} | \mathrm{w}_s, \vec{r}_s, b)\,, \tag{5.5}$$

where the i^{th} instance of Eq. 2.18 from a source at position $\vec{r}_s$ is given by

$$\bar{x}_i = v_{is}\mathrm{w}_s + bt_i\,, \tag{5.6}$$

and the system response for measurement i from a point-source at $\vec{r}_s$ (neglecting attenuation) is

$$v_{is} \approx \frac{\eta(\vec{r}_s, \vec{r}_i) t_i}{|\vec{r}_i - \vec{r}_s|^2}, \tag{5.7}$$

where $\eta(\vec{r}_s, \vec{r}_i)$ is the angular response of the detector at position $\vec{r}_i$ to a point-source at position $\vec{r}_s$. For an isotropic detector η = constant.

In this formulation, the response is calculated only where needed in the optimization. This removes the need to compute the entire 3D system matrix, $\mathbf{V}$, significantly reducing the computational and memory burden of the reconstruction. Furthermore, the reconstruction is no longer limited to the size of voxels used to discretize the image space.

Note that solving for source intensities and source positions are independently convex problems, but solving for them simultaneously in Eq. 5.5 is no longer convex. Several methods exist to overcome this issue, but for this work a traditional convex optimization algorithm was used along the optimal ML-EM intensity (and background) manifold. In other words, the optimization algorithm was allowed to search only the positions of the point sources and at each step (source position) ML-EM was performed to find the optimal source intensity and constant background.

The Bound Optimization BY Quadratic Approximation (BOBYQA) derivative-free algorithm [144] for bound constrained optimization in the Non-Linear Optimization (NLOPT) library [145] package within the PYthon parallel Global Multiobjective Optimizer (PYGMO) library [146] was used here with 20 ML-EM iterations at each step. Several other algorithms (both derivative-free and gradient-based) are available, but BOBYQA produced the best results with the fastest convergence in this problem. The results are shown in Fig. 5.7. The point-source position is no longer limited by the voxel size, producing a position error < 2 cm, and the error in the reconstructed point-source intensity was $< 1\%$ of the true intensity.

The reconstruction runtime was 160 ms, considerably faster than the previous approaches. This time also includes the calculation of the system matrix, but in this case the response was only computed for $\sim$50 positions, equivalent to $\sim$50 columns of $\mathbf{V}$. This is compared to the $J = 1.28 \times 10^5$ columns for the voxelized space in the previous section. The 50 columns in this case indicate the number of positions tested in the BOBYQA optimization routine before convergence. Note that as the number of measurements and voxels increases, the full calculation of $\mathbf{V}$ becomes significantly expensive and memory intensive (in some cases, requiring more memory than available RAM, thus further slowing the reconstruction as $\mathbf{V}$ must be computed on the fly), further limiting the runtime performance of spatially-discretized approaches. The continuous PSL formulation avoids this problem entirely, scaling linearly with the number of measurements.

To explore the performance of the continuous PSL approach in this source scenario,

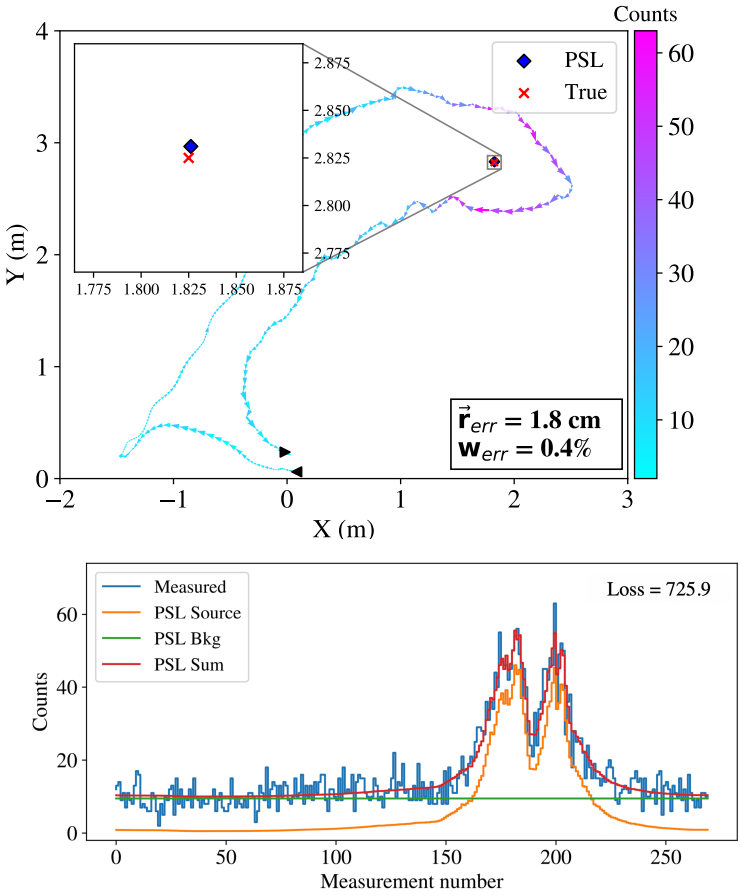

Figure 5.7. (Top) PSL solution (blue diamond) of a single point-source to best describe the measured data, solved for in continuous 3D space along the maximum likelihood source intensity and background manifold using a conventional derivative-free optimization algorithm (BOBYQA). (Bottom) Forward projected mean-rates $\hat{\boldsymbol{x}}$ compared to the measurement $\boldsymbol{x}$.

the simulation was run 10,000 times with source activities randomly sampled between $1-10$ μCi and a constant background rate of 100 counts per second. Figure 5.8 shows a box plot of the position and source intensity errors against the true source strength, binned at 1 μCi intervals. For each interval, the mean signal-to-noise-ratio ($\overline{\text{SNR}}$) is shown as an additional horizontal axis. The SNR of a single simulation is defined as $\max(S_i/\sqrt{S_i+B_i})$ where S_i and B_i are the signal and background at measurement i, respectively. The position errors are expressed as a percent of closest approach (45 cm). The median in each bin is denoted with an orange horizontal line, the boxes range from the first to third quartile, and the arms extend out to the 10^{th} and 90^{th} percentiles. Both the median and spread of each error decreases with source strength, as expected. The median source intensity error is below 10% across all activities and the median position error is less than 40% of the distance of closest approach.

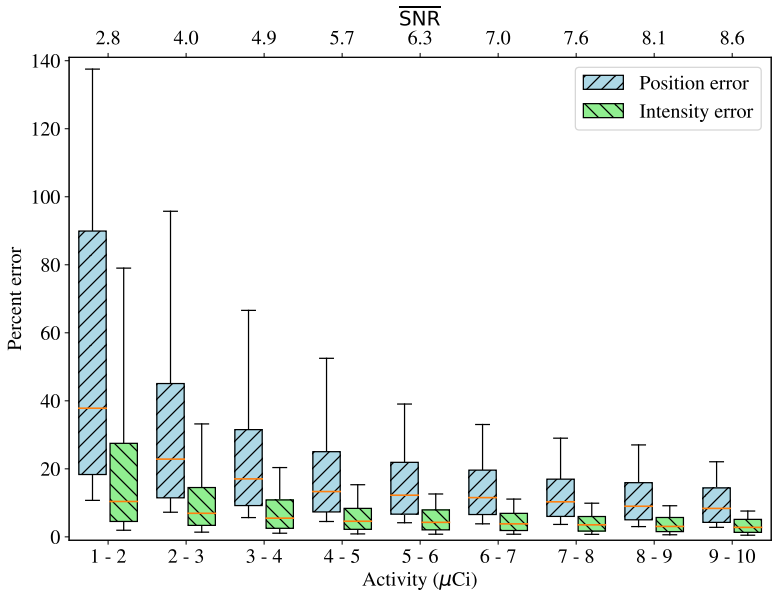

Figure 5.8. Box plot of the position and source intensity errors against the true source strength, binned at each μCi, for 10^4 simulations with source intensities randomly sample between $1-10$ μCi. Position error as defined in Sec. 2.5.3 is shown here in percent of closet approach. The median in each bin is denoted with an orange line, the boxes range from the first to third quartile, and the arms extend out to the 10^{th} and 90^{th} percentiles. The mean SNR over each interval is shown as an additional horizontal axis.

Figure 5.9 shows the correlation between the position and source intensity errors for $>$ 99% of the total 10^4 simulations (outlier reconstructions are ignored). Recall the detector position remained close to $Z \approx -14$ cm while near the source, which was placed at $Z = 0$. The isotropic response of the detector and the little movement in Z results in degenerate solutions in likelihood that extend along Z. Furthermore, the position below the measurement path equal to the offset of the path to the true source location ($Z \approx -28$ cm or 60% position error) represents a degenerate solution in source intensity. These effects can be seen in the plot in three distinct regions (shown with dashed green lines). First, a diagonal band extends out from the origin, corresponding to reconstructed positions above the true location in Z. A second diagonal band extends out from a position error of 60%, corresponding to reconstructed positions below the degenerate source intensity solution in Z. Finally, a slightly curved band is observed between the origin and this point, representing the positions between the true source location and the degenerate source intensity solution on the opposite side of the path. The band peaks at a position error of 30% (plane of the measurement path, $Z \approx -14$ cm), where the source intensity estimate is lower to compensate for the closer source. Ultimately, while some degeneracy is observed in the solution space, enough variation existed in the detector position that the majority of reconstructions were near the origin in Fig. 5.9, correctly

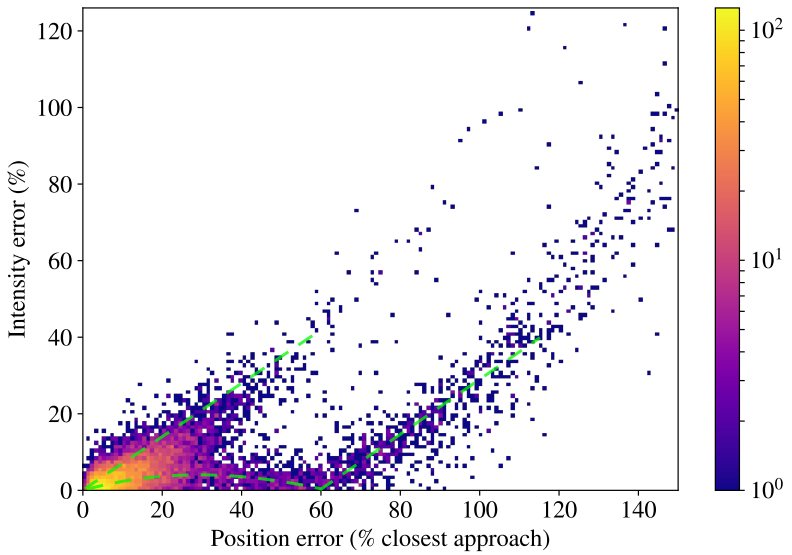

Figure 5.9. 2D histogram of position and source intensity errors for $> 99\%$ of the total 10^4 simulations with source activities between $1-10$ μCi. Less than 1% of events were outside of the defined bounds and are not shown. Bin widths are 1% in each direction. The degeneracy of the solution space is highlighted by green dashed lines and a discussion in provided in the text.

localizing and quantifying the source, particularly for high SNR sources.

5.5 Additive Point Source Localization

The additive nature of Poisson variables facilitates the inclusion of M known, constant source rate contributions, $\boldsymbol{\mu}^{[I\times 1]}$ into Eq. 5.6, rewritten as

$$\bar{x}_i = v_{is}\mathrm{w}_s + \mu_i + bt_i\,, \tag{5.8}$$

where $\mu_i = \sum_{m=1}^{M} v_{im}\mathrm{w}_m$. Now Eq. 5.5 can be reformulated to localize an additional source

$$\underset{(\mathrm{w}_s,\vec{r}_s,b)}{\operatorname{argmin}}\ \ell(\boldsymbol{x}|\mathrm{w}_s,\vec{r}_s,b,\boldsymbol{\mu})\,. \tag{5.9}$$

While one can attempt to solve Eq. 5.9 in discrete space as in Sec. 5.4.1 by successively solving for sources one at a time, this approach does not allow for the re-optimization of source intensities and positions after a new source is found. In the continuous space formulation, a re-optimization can be done in position and intensity across multiple sources, allowing all the free parameters to vary at once. This capability is crucial in the multi-point-source reconstruction problem as the addition of another source can affect the overall likelihood fit of the previous source configuration.

Algorithm 1 Additive Point-Source Localization

```
Initialize reconstruction. S = {(); b = median(x)}
converged = False
Solve Eq. 5.5, append to S
while not converged do
    S_old = S
    Solve Eq. 5.9 for additional source, append to S
    Re-optimize source positions, intensities and background, update S
    Test for acceptance of S relative to S_old using BIC
    if accepted then
        Clean S: drop low weight or weakly contributing sources and collapse nearby sources
        Re-optimize current state of source positions, intensities and background, update S
    else
        S = S_old
        converged = True
```

Algorithm 1 is proposed here to iteratively reconstruct a sparse parametric image of N sources, $\mathbf{S} = \{(\mathrm{w}_1, \vec{r_1}), ..., (\mathrm{w}_N, \vec{r_N}); b\}$, in the continuous space formulation, where N is also treated as an unknown. After each new source is identified, the re-optimization of source positions and intensities can be done in two ways:

1. In the fashion of [147], alternate between
    a) Fix intensities and re-optimize positions using conventional optimization methods (e.g., BOBYQA).
    b) Fix positions and re-optimize intensities and backgrounds using ML-EM.
2. Re-optimize source positions using conventional optimization methods along the optimal intensity and background manifold (as done in Sec. 5.4.2).

A model selection criterion is then used to test the new model ($N + 1$ sources) compared to the old model (N sources). The Bayesian information criterion (BIC) [148] was used here and is given by

$$\mathrm{BIC} = \log(I)k + 2\ell(\boldsymbol{x}|\hat{\hat{\boldsymbol{x}}}), \tag{5.10}$$

where $k = 4N + 1$ is the number of parameters estimated by the model (the intensity and XYZ position of each source and the constant background rate in the detector) and I is the number of measurements. Again, for a multi-detector system k would generalize to $k = 4N + N_\mathrm{d}$. The preferred model is the one with the minimum BIC value. The BIC penalizes the acceptance of a new model based on the number of parameters used in the model (i.e., the addition of another source must significantly improve the model in order to be accepted). The BIC was used here, as opposed to another model selection criterion such as the Akaike information criterion (AIC) [149], as it applies a stricter penalty, therefore restricting the addition of many sources and reducing overfitting.

If the new model is accepted, a cleaning procedure is invoked to drop low intensity and weakly contributing sources as well as combine spatially close sources (defined to be < 10 cm in this case). If the source configuration is changed in the cleaning, a re-optimization is performed again. If the new model is rejected, the algorithm is stopped. APSL correctly stopped at one source in the source scenario presented above.

The multi-source performance of the APSL algorithm is explored using the simulated detector and measurement path from above (see Fig. 5.10). In addition to the 5 μCi source used previously, three additional sources were placed in the XY plane near the trajectory, with activities ranging from $6-8$ μCi and closest standoffs of $35-55$ cm. The bottom pane of of Fig. 5.10 shows the simulated signal source components and background at each measurement.

The unregularized ML-EM reconstruction, in $\log_{10}$ space, along the XY plane and the corresponding fit in count space are shown in Fig. 5.11. Forty iterations were required in this case to achieve the same likelihood convergence stated in Sec. 2.5.3. The system matrix calculation runtime was 22 s and the reconstruction runtime was 1.2 s. ML-EM continues to produces an overfit solution and fails to correctly localize any of the sources. Again, the overfitting behavior worsens with more iterations. Similar to Sec. 2.5.4, the inclusion of sparsity regularizers in ML-EM did not improve the results (see Figs. 5.12 and 5.13).

The APSL reconstruction results are shown in Fig. 5.14 and the errors are shown in Table 5.1. The BOBYQA algorithm was used with 20 ML-EM iterations at each step in the position optimization (additional iterations were not necessary and did not affect the result). The total reconstruction runtime was 12.8 s. Again, this time includes the calculation of the columns in $\mathbf{V}$ needed in the optimization. APSL converges to the correct number, location, and intensity of the four unique sources, without the use of pre-conditioning as in [132], and accurately estimates the constant background rate.

The reconstruction errors in position are all < 10 cm in the XY plane, with slightly higher errors in the Z dimension. This effect is more pronounced for the sources outside of the measurement path (#1 and #4). This is expected again due to the degeneracy in the solution space discussed previously. In general, it is shown that larger position errors correlate with larger intensity errors, as was shown in Fig. 5.9.

While this example uses a simple model with nearby sources and relatively large individual SNRs, several peaks overlap to produce an overall complex source term. Furthermore, the measured detector path captures the variation in detector position expected in an operational source search scenario, adding variation to the signal as well as breaking some of the degeneracy in the ML solution space. Ultimately, the results highlight the success of the APSL algorithm in the deconvolution of the signal into the correct individual source components and background.

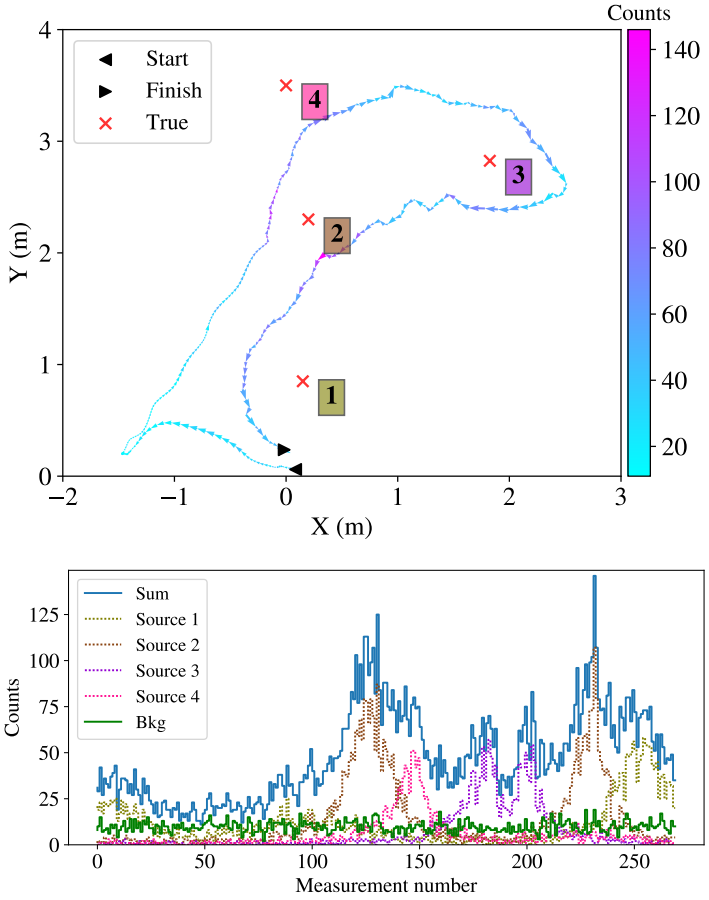

Figure 5.10. (Top) Measurement scenario similar to Fig. 5.1 now with multiple sources in the XY plane. The source positions are shown with red x's as well as numbers in colored boxes. (Bottom) Simulated signal source components and background at each measurement.

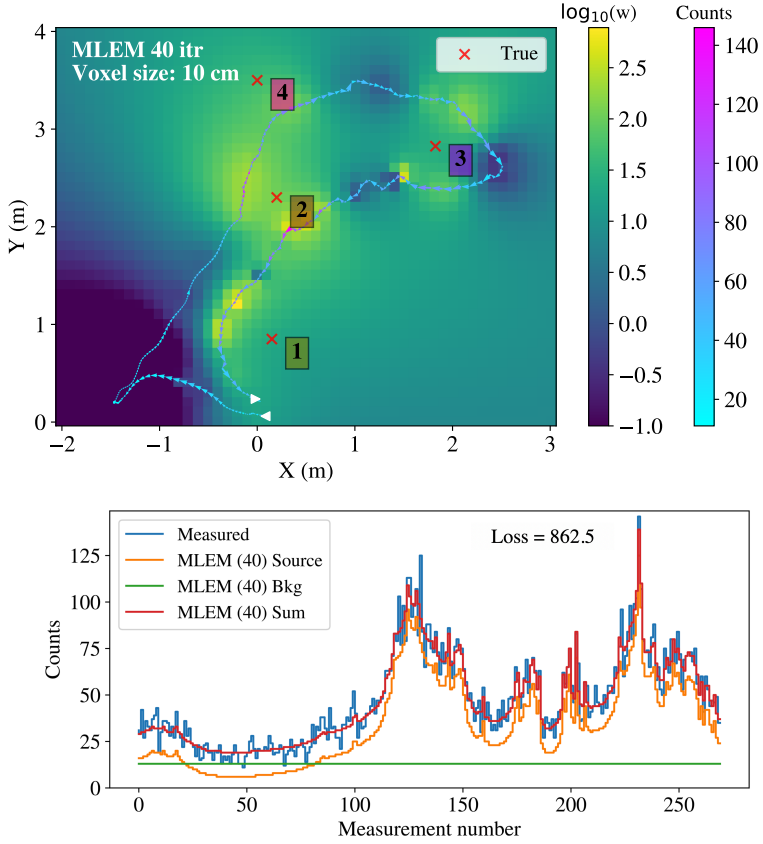

Figure 5.11. (Top) ML-EM reconstruction $\hat{\mathbf{w}}$ (40 iterations), in $\log_{10}$ scale, of a slice through the XY plane (Z = 0) for the scenario presented in Fig. 5.10. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{x}}$ compared to the measurement $\boldsymbol{x}$.

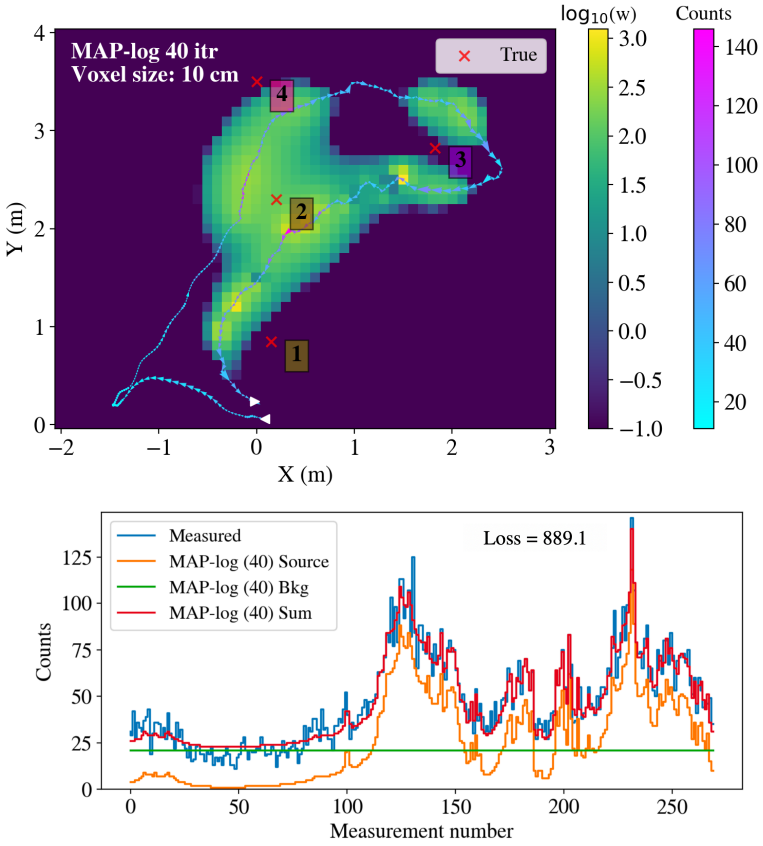

Figure 5.12. (Top) MAP reconstruction $\hat{\mathbf{w}}$ (40 iterations) with log prior, in $\log_{10}$ scale, of a slice through the XY plane (Z = 0) for the scenario presented in Fig. 5.10. (Bottom) Forward projected mean-rates $\hat{\hat{x}}$ compared to the measurement x.

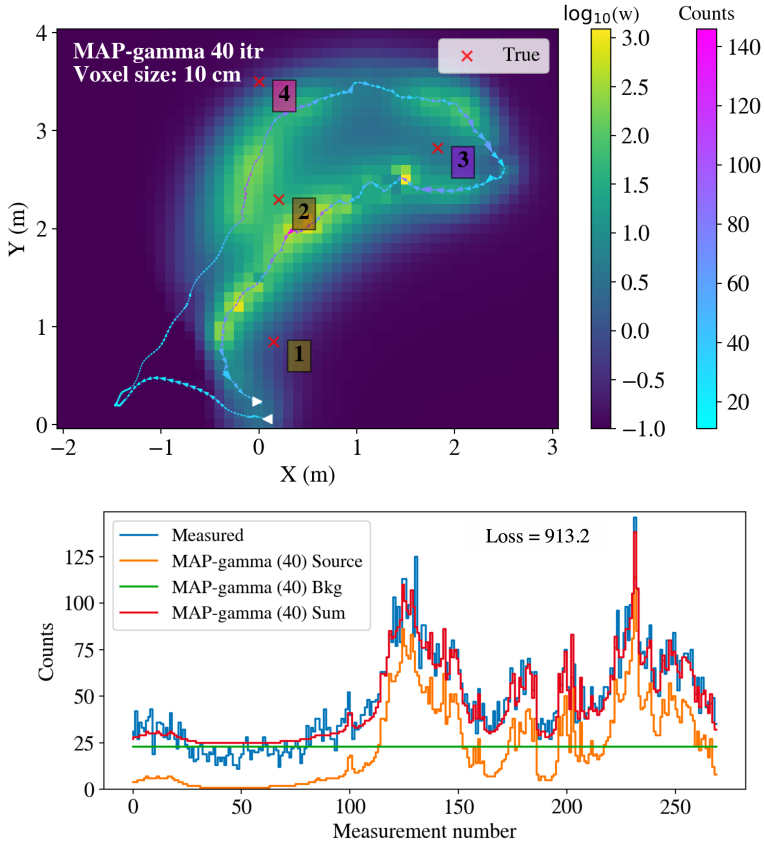

Figure 5.13. (Top) MAP reconstruction $\hat{\mathbf{w}}$ (40 iterations) with gamma prior, in $\log_{10}$ scale, of a slice through the XY plane (Z = 0) for the scenario presented in Fig. 5.10. (Bottom) Forward projected mean-rates compared to the measurement.

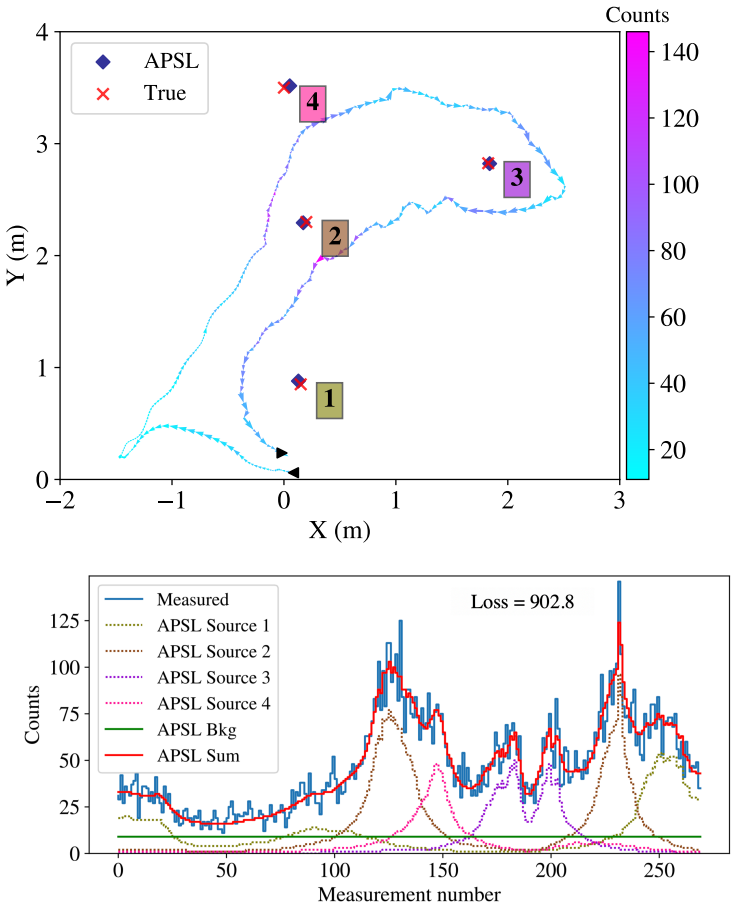

Figure 5.14. (Top) APSL solution (blue diamonds) in the XY plane (Z = 0) for the scenario presented in Fig. 5.10. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{x}}$ (individual source and background components) compared to the measurement $\boldsymbol{x}$.

Table 5.1. APSL reconstruction errors in Fig. 5.14

	Source 1	Source 2	Source 3	Source 4
$\mathbf{w}_{\text{true}}$ (μCi)	8.0	6.0	5.0	6.0
$\hat{\mathbf{w}}$ (μCi / % err)	7.3 / 9.9	5.6 / 7.2	4.8 / 3.9	7.7 / 22.3
$\hat{\hat{r}}_{\text{err,XY}}$ (cm)	3.6	2.9	1.3	5.5
$\hat{\hat{r}}_{\text{err,Z}}$ (cm)[1]	10.5	5.8	5.4	15.4

[1]Generally, $\hat{\hat{r}}_{\text{err,Z}} > \hat{\hat{r}}_{\text{err,XY}}$ due to the non-directional detector used and that the path primarily moved in XY, without much change in Z.

5.5.1 Separation of Spatially Close Sources

The ability to separate sources with an isotropic detector is dependent on the measurement path, the statistics collected, and the degeneracy of the ML solution space. However, it is expected that, given a path near two sources, an isotropic detector should be capable of resolving two sources spatially separated by at least the distance of closest approach.

To demonstrate the capability of APSL to resolve spatial close point-sources, two unequal activity sources (9 μCi and 5 μCi) are simulated in the XY plane near the measurement path, with a 40 cm distance to closest approach. Two scenarios in which the source separations are 40 cm and 30 cm are considered. APSL is run on each scenario and the results are shown in Figure 5.15.

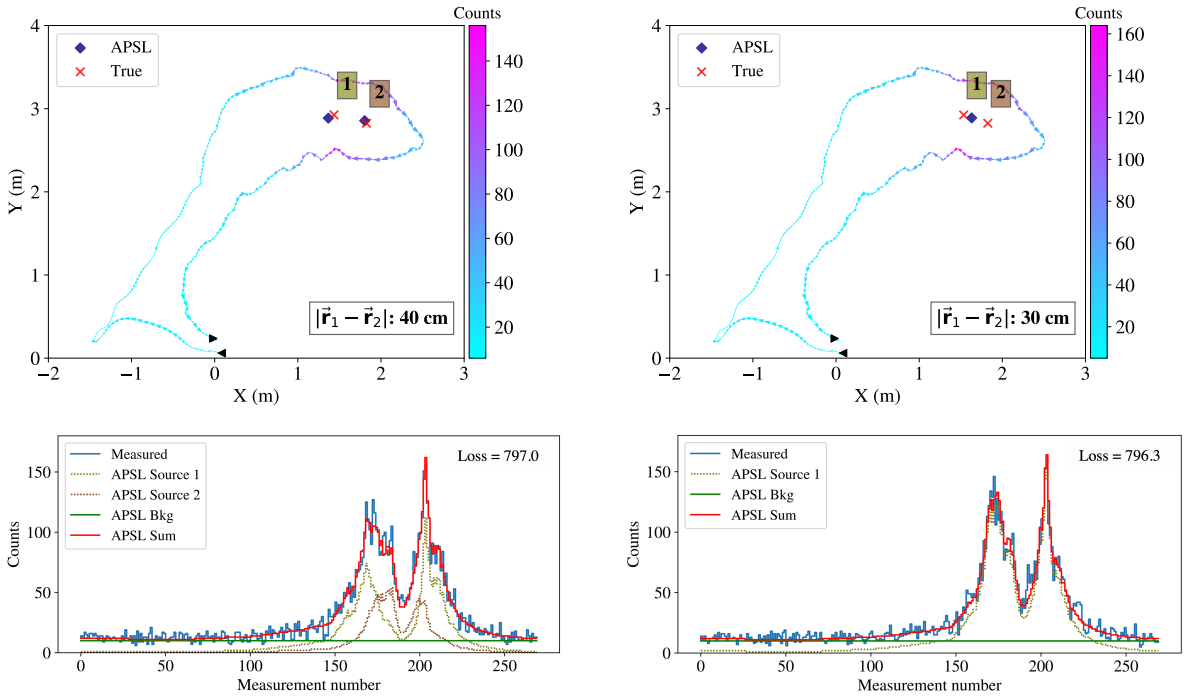

Figure 5.15. (Top) APSL solution (blue diamonds) in the XY plane (Z = 0) for two sources (#1: 9 μCi, #2: 5 μCi) with separation of 40 cm (left) and 30 cm (right) placed in the XY plane. The distance of closest approach in the measurement path is 40 cm in both cases. (Bottom) Forward projected mean-rates $\hat{\boldsymbol{x}}$ (individual source and background components) compared to the measurement $\boldsymbol{x}$ for the images in the top panel. Reconstruction errors are discussed in the text.

In the case where the source separation is the distance of closest approach, APSL is successful in resolving two sources, with XY position errors < 10 cm, Z position errors < 15 cm, and intensity errors $< 5\%$. In the case where the source separation is less than

the distance of closest approach, APSL places a single source with a combined activity (13.8 μCi) between the two sources. A slight bias is observed towards the stronger source (#1), though this is expected as this source contributes more to the measured signal and a better fit to this source can result in a lower overall loss.

5.5.2 Experimental Demonstration

Here the experimental performance of APSL is explored. Data was collected with a hand-held free-moving 2 in (5.08 cm) right-cylinder $LaBr_3$ scintillator with a SDF contextual sensor package (see Fig. 5.16) similar to the PRISM system (LiDAR, IMU, RGB camera, board computer). The system was assumed to have uniform directional sensitivity for reconstruction purposes, ignoring the occlusions from the contextual sensors and any directionality in the $LaBr_3$ crystal itself. SLAM was performed with Google Cartographer to produce a 3D map and pose estimates at $\sim$10 Hz as the system was walked around the scene. Here a comparison is made between APSL and SDF-constrained ML-EM. It is important to note here that the constant background assumption is actually quite strong with this system, due to the high intrinsic radiation background due to the naturally occurring ^{138}La radioisotope (10^{11} yr half-life) in lanthanum (producing 789 and 1436 keV gamma-rays) [150].

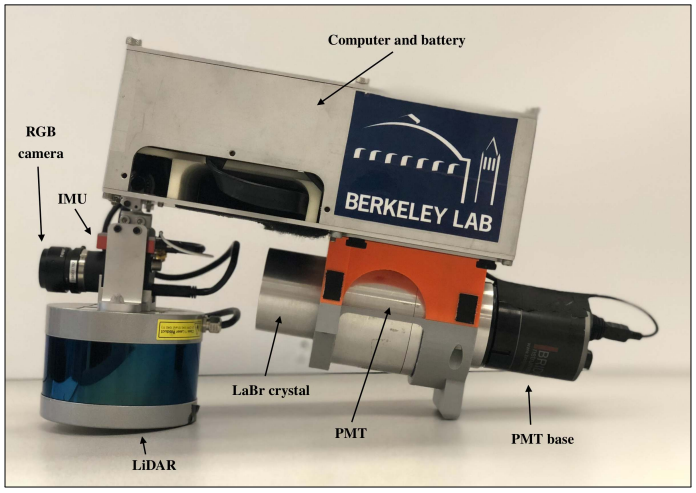

Figure 5.16. Picture of a free-moving 2 in (5.08 cm) right-cylindrical hand-held $LaBr_3$ spectrometer with a SDF contextual sensor package (LiDAR, IMU, RGB camera, board computer).

The measurement scenario considered contained three well-separated ^{137}Cs sources of activities 5, 10, and 20 μCi placed at varying elevations in a small ($\sim$100 m^2) indoor cluttered laboratory space. The system was walked slowly around the perimeter of the laboratory space, moving the system up and down in an attempt to break solution degeneracy in elevation. The total measurement time was 2.08 min (1240 poses). Figure 5.17

shows a top down projection of the final 3D point cloud with yellow starts denoting the true source locations in the scene and the measurement path colorized by the measured gross count-rate. The beginning and end of the measurement are shown with black right-facing and left-facing triangles, respectively. The count-rate is also shown in the lower pane as a function of the measurement number. Gross-counts were used here as a general approach (i.e., no priors on the source energies), though the spectroscopic capability of $LaBr_3$ facilitates radioisotope identification and reconstructions could, in prinicple, be run on specific ROIs.

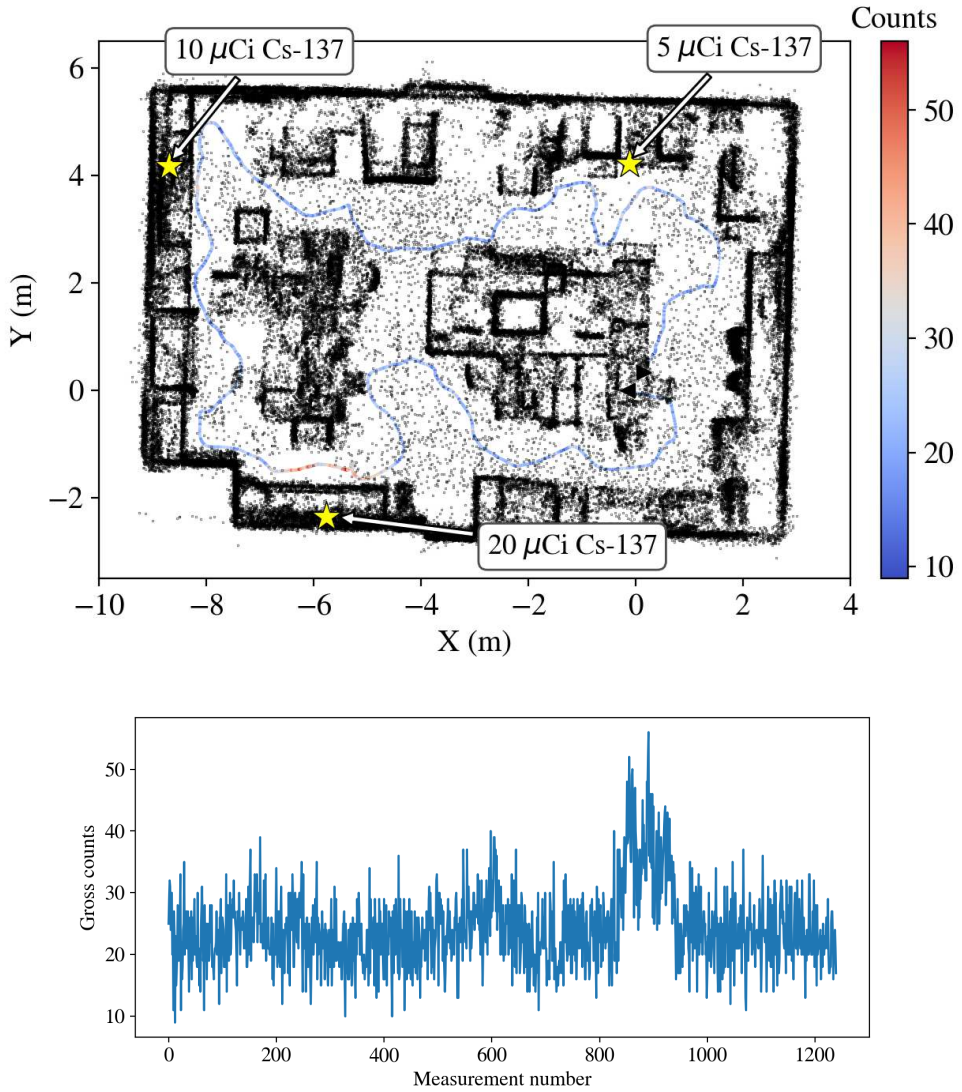

Figure 5.17. (Top) Top-down projection of the experimental measurement path and point cloud. The path is colorized by the gross count-rate and (bottom) shown as a function of the pose index.

The ML-EM reconstruction (40 iterations) is shown in Fig. 5.18 as a colorized point cloud with the lower 10% of the gamma-ray intensities replaced with the LiDAR return intensities for added contrast. The image space was uniformly voxelized with 15 cm voxels over the entire extent of the point cloud (4.5×10^5 total voxels). An occupancy criteria of 10 points and 4 neighbors was applied and used to constrain the reconstruction (4.3×10^4 occupied voxels). The system matrix and image reconstruction compute time was 35 s on the CPU hardware described above. The high intensity regions indicate that the ML-EM solution is successful in coarsely localizing regions of the three sources. The SDF-constraint is seen to significantly improve the results (relative to the unconstrained reconstructions in previous sections), however source biasing and overfitting is still observed. Similar to previous results, ML-EM neglects the background parameter (estimating $\hat{b} = 0$) and places source intensity everywhere in the image space to compensate.

Figure 5.19 shows the APSL reconstruction, with green squares denoting the location of the reconstructed sources. The reconstruction time on the CPU hardware used above was 5 s, showing significant improvement over ML-EM. APSL is successful in determining the number of sources in the measurement (3) and with accurate localization (within tens of cm) in 3D. There is some slight biasing towards the measurement path, due most likely to the degeneracy in the solution space and insufficient spatial sampling and sensitivity. The system was continuously moved up and down (in elevation) during the measurement in an attempt to break the degeneracy in the Z dimension and the results show that this approach was successful. However, no significant effort was made to the move the system closer and farther away from potential source locations and thus, with a single pass-by of the sources, left some degeneracy in the radial dimension outward from the track. This slight bias towards the track also accounts for the inability of APSL to accurately discriminate the source intensities of the 5 and 10 μCi sources. A much more accurate constant background rate estimate is obtained with APSL. The kinks and notches in the background component arise due to nonuniform time durations of the measured poses (see Fig. 5.20).

Due to developments in ML-EM acceleration described in previous chapters, it is important to state that the reconstruction took 7.4 s with a GPU-accelerated bin-mode SDF-ML-EM implementation on a Radeon Pro 455 dedicated GPU. A bin-mode formulation was used here because the number of poses (1240) was significantly less than the total number of counts (3×10^4). The difference here compared to the previous chapters with PRISM is attributed to the use of a single detector and gross counts. The CPU-based APSL formulation still outperforms the GPU-accelerated SDF-ML-EM reconstruction. Future work could be done to further optimize APSL for computational efficiency and GPU-acceleration.

An isotropic response was assumed here in both the APSL and ML-EM reconstructions. Preliminary simulations have shown that while the crystal itself has a small vari-

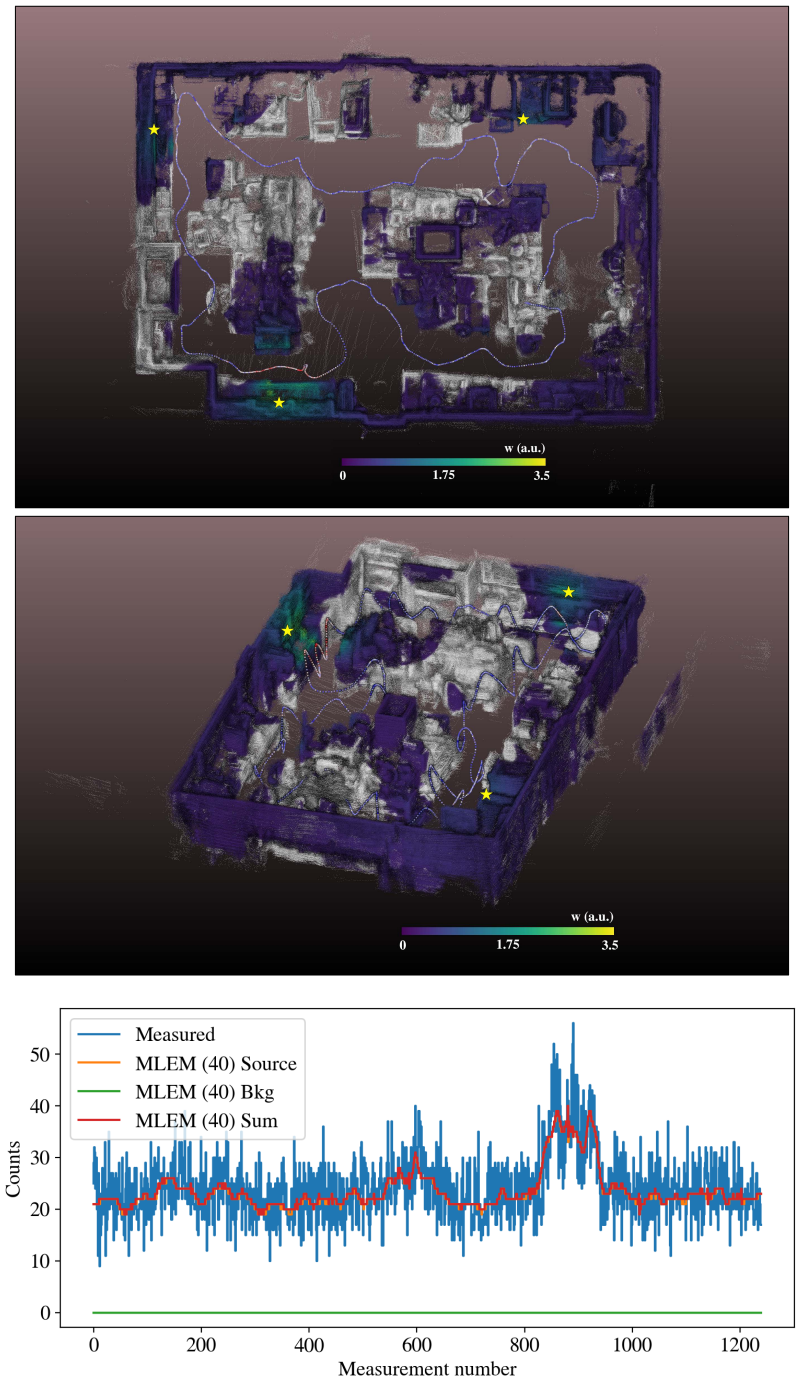

Figure 5.18. (Top) SDF-constrained ML-EM reconstruction (40 iterations) shown as a colorized 3D point-cloud with the measurement path colorized with the measured gross count-rate. True source locations are shown with yellow stars. (Bottom) Forward projected mean-rates compared to the measurement.

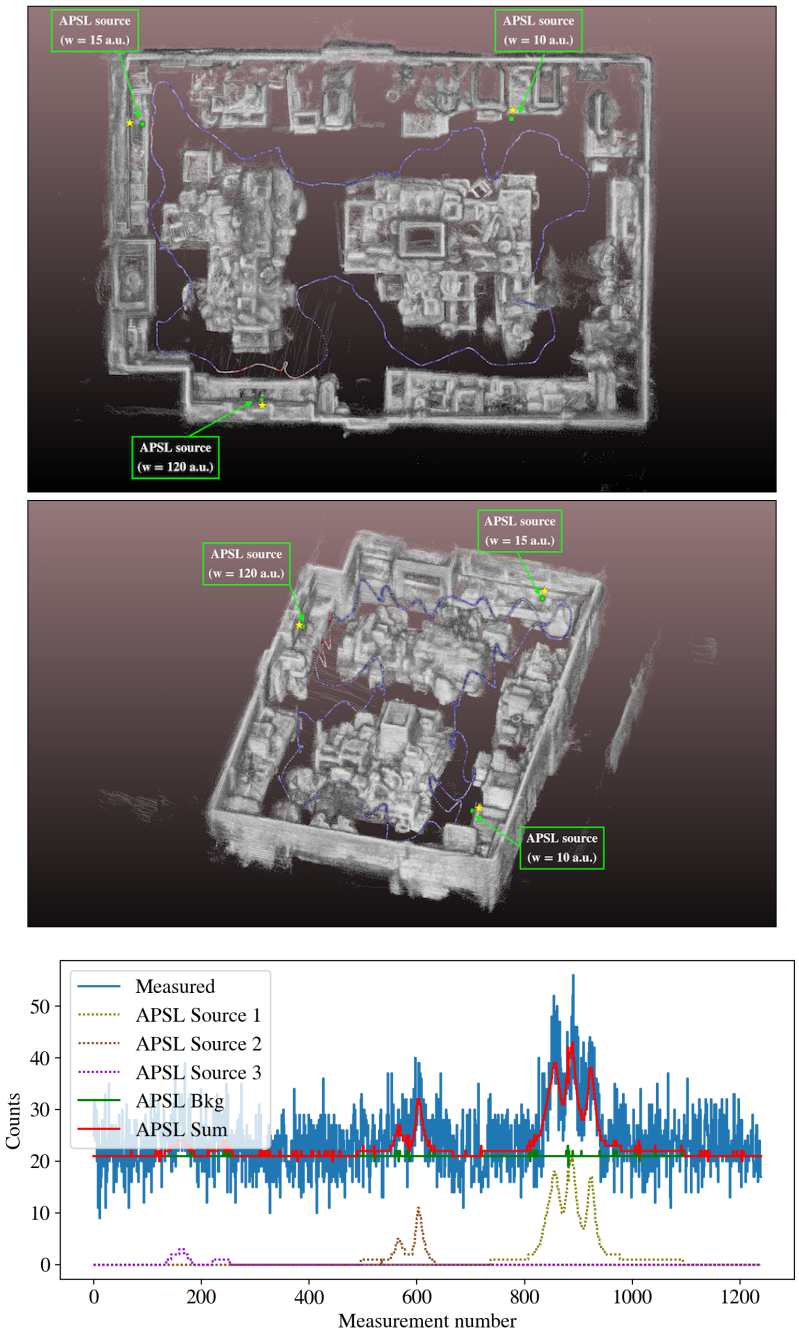

Figure 5.19. (Top) APSL reconstruction shown as green squares superimposed on the 3D point-cloud with the measurement path colorized with the measured gross count-rate. Reconstruction source activities are shown in arbitrary units as the system response used was not quantitative. True source locations are shown with yellow stars. (Bottom) Forward projected mean-rates compared to the measurement.

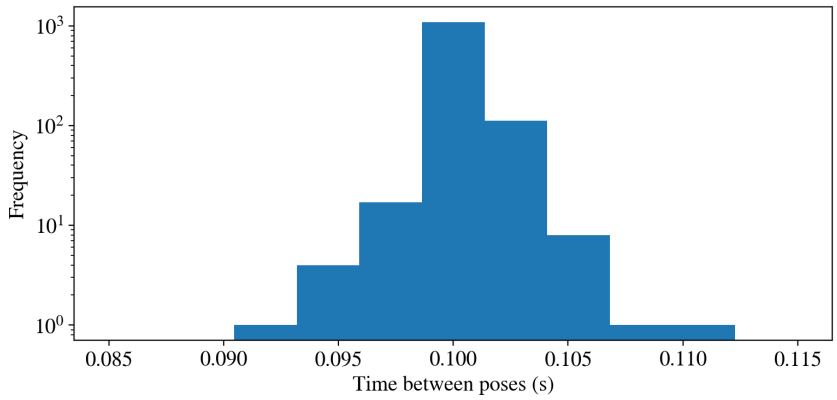

Figure 5.20. Distribution of time between poses collected with a contextual sensor package for the APSL experimental measurement.

ation in directional sensitivity (∼5% across the polar angle at 662 keV), the LiDAR and battery act as significant attenuators to the $LaBr_3$ response, even at 662 keV. These features in the directional response may aid in the breaking of degeneracy and improve source location. The implementation of directional information into the APSL framework will be the subject of future work.

5.6 Conclusions and Future Work

Using a simple simulation of a single point-source and a non-directional detector moved over an experimental SLAM-tracked trajectory, it has been shown that a reformulation of the sparse inverse problem can improve source localization and quantification performance. In sparse scenarios, the traditional and sparsity-enforcing regularized ML-EM approaches suffer from overfitting and fail to localize point-sources. Ultimately, these approaches are attempting to solve an inherently under-determined problem. The reformulation of the problem with a point-source assumption (PSL) is well-posed and was shown to accurately localize point sources and quantify their intensity, though at the expense of computational complexity when reconstructing in a discretized spatial domain. Solving PSL as an optimization over continuous intensity and position further increased localization and quantification performance and improved computational efficiency and memory consumption.

The continuous PSL formulation was extended to APSL for multiple point-source localization or sparse parametric image reconstruction. To demonstrate the concept, four point-sources were simulated near the same experimental measurement path, forming a complex source term. ML-EM failed to localize any of the sources or provide a useful initialization for a gradient-based approach, even with sparsity regularization. APSL

successfully deconvolved the signal into the correct individual source and background components (including the number, location and strength of the sources) at a reduced computational and memory burden.

APSL also successfully resolved (in both space and intensity) two simulated sources of unequal activity in close proximity. The resolving capability drops quickly near the distance of closest approach, but detectors or detector arrays that are better able to determine the direction of photon incidence are expected to better resolve closer sources.

Finally, APSL was demonstrated experimentally with a free-moving $LaBr_3$ detector with a SDF contextual sensor package and compared to the SDF-constrained ML-EM approach. ML-EM was able to broadly localize three well-separated sources, however still suffered from source biasing and overfitting. APSL converged to the correct number of sources and accurately localized the sources in 3D to within tens of cm. APSL was able to discriminate the source activities, though only on a relative scale because the system matrix used in the reconstruction was not quantitative.

Further work is needed to rigorously characterize the algorithm and validate the scenarios in which it may be used. This includes investigations into complex source terms such as weak, large standoff, and shielded sources as well as more constrained measurement paths relevant to radiological source search (e.g., straight lines or single pass) that may be subject to more degeneracy in the solution space. The impact of imaging detector systems and/or arrays (i.e., anisotropic detectors) on localization accuracy and computational burden should be explored. Another possible route of investigation could be into the use of APSL for extended or compact non-point-like sources, adding an additional free parameter describing the shape of the source (e.g., a radius for a finite size sphere or a length for a line source).

While a constant background assumption was made here and may be appropriate for the small search space and short duration of the measurements, this may not be appropriate in wide-area urban search scenarios [151]–[153]. Therefore a treatment for variable background rates must be incorporated into the algorithm for large search spaces and longer measurements. For example, advanced radiological detection and identification algorithms designed for dynamic background environments [82] could be used to model spatially or temporally-varying background counts in the measured signal. The counts attributed only to an identified source (or sources) could then be used as an input to the APSL framework, mitigating the need for a varying background term. This approach would also facilitate the localization of specific radioisotopes.

Finally, relevant for field operations, APSL could be developed for streaming operation, informing the user of potential source locations in real-time and directing the path to decrease the time to detect and localize sources. This could include constraining the search space to static surfaces and volumes (i.e., SDF) or to moving objects in the 3D scene model (i.e., object tracking [154]–[156]), improving the computational speed of the reconstruction as well as providing additional context to the user in real-time.

Chapter 6

Conclusions

6.1 Summary

Gamma-ray imaging has relevant applications in many fields ranging from astrophysics and medical imaging, to nuclear facility decommissioning and proliferation detection. In the application space presented here, namely nuclear security, there exists a critical need for the capabilities to efficiently localize compact sources of radiation as well as to image or map extended source distributions in small cluttered spaces and large urban environments. This work set out to enhance the current set of state-of-the-art through the development and demonstration of novel radiation detection instrumentation and free-moving 3D imaging techniques for low-energy (< 450 keV) gamma-rays.

First, a spherical active coded mask imaging system was designed to resolve the limited coded FOV and non-uniform sensitivity issues inherent in traditional planar coded mask imaging systems. The PRISM prototype, a hand-held, omnidirectional, broad energy sensitive, multi-modal CPG-CZT-based imager, was assembled, characterized and successfully demonstrated. The coded mask was optimized to maximize sensitivity, localization performance, and uniformity over 4π, using a Great Deluge optimization framework and a fast far-field zero-energy system response generator. Equal weightings were given to the three optimization metrics in order to design a balanced system with regards to detection efficiency and imaging performance. In the future, different weightings or entirely new metrics could easily be implemented into the optimization framework to tailor the design for specific applications as they arise.

In contrast to static imagers, free-moving systems facilitate efficient source localization with the ability to overcome the inverse square law and can circumvent potential shielding scenarios by quickly viewing sources from multiple perspectives. Moreover, the integration of SDF into these systems has enabled real-time 3D mapping in both the contextual and gamma-ray domains. This work represents the first experimental demonstration of an approach to low-energy real-time 3D gamma-ray imaging with SDF, using

PRISM and a GPU-accelerated list-mode ML-EM reconstruction approach. Source localization was successful over energies from $60-356$ keV, activities from $20-5000$ μCi, and areas of size $70-4500$ m^2. Indoor and outdoor scenarios were considered and measurements were all < 3 min. Reconstruction was performed in both the near- and far-field regimes and image quality improvements were demonstrated with DOI read-outs.

A novel PSL algorithm was developed and successfully demonstrated to mitigate the limitations inherent in traditional grid-based ML-EM reconstruction approaches in sparse scenarios, namely overfitting and computational complexity. The PSL formalism was extended to APSL for multiple point-source localization and general sparse parametric image reconstruction. APSL successfully reconstructed multiple sources in 3D using a non-directional free-moving detector, in both simulation and experiment, with better accuracy and reduced computational burden than ML-EM. Continued work is necessary to characterize and optimize the algorithm for various scenarios. This includes more difficult source terms such as weak, shielded, and non-point-like sources and well as the additional gain from directional detector responses and detector arrays.

6.2 Future Outlook

The results presented in this work represent a subset of the possible applications of the SDF concept. While a hand-held system using the active coded mask imaging modality was used here for demonstration, the approach is generally agnostic to the detection platform (i.e., hand-held, ground-vehicle, aerial-borne), the operation (i.e., manned, unmanned), the imaging modality (i.e., proximity, coded mask, Compton), the detection media (i.e., gas, semiconductor, scintillator) and the radiation signature (i.e., gamma-rays, neutrons, etc.). This work was limited to imaging unshielded gamma-ray point-sources, but the approach can be used to image shielded, extended, and distributed sources. This has been demonstrated at higher energies using Compton and proximity imaging [9], [29], but additional effort is necessary to demonstrate this in the low-energy regime.

While hand-held systems can be used for small-scale search and mapping on the ground, UASs enable efficient wide-area mapping and remove the need to expose human operators to potentially high dose-rate environments. Progress towards autonomous navigation and multi-platform (e.g., UAS swarms) mapping should be continued for large-scale SDF. Small UAS operation is limited by payload, thus continued effort is required to design and fabricate lighter, more compact, and lower power SDF capable systems. The PRISM sphere is optimal for 4π imaging and sensitivity, however, the design presents a significant engineering challenge and cannot easily be scaled to smaller sizes without a loss in detector mass (i.e., sensitivity) as the modules are currently packed as tightly as possible. More compact and scalable configurations are currently being

explored at LBNL, for example a 6-plane active coded hollow cube and a fully 3D cube using stacked coded planes.

Different detection media are also being explored at LBNL, particularly new high-density inorganic scintillator materials, specifically elpasolites such as Cs_2LiYCl_6:Ce (CLYC) and $Cs_2LiLaBr_{6-x}Cl_x$:Ce (CLLBC), that produce high light-yield, have excellent energy resolution ($3-4\%$ at 662 keV), and are neutron sensitive and capable of pulse shape discrimination (PSD) [157]–[159]. These crystals can be grown to larger sizes to accommodate more detector mass in smaller overall form factors and can be coupled to multichannel read-out systems for 3D position sensitivity.

Advanced gamma-ray imaging systems facilitate high resolution mapping necessary for certain applications such as, for example, hold up characterization in nuclear facilities [160], [161]. In the absence of sophisticated imaging systems or in scenarios where high-resolution maps may not be necessary, the SDF formalism developed in this work can be applied directly to proximity imaging and mapping systems that may be cheaper and more scalable. Furthermore, in the case of sparse radiological source search, this work has shown that APSL can produce accurate source localization estimates of multiple sources in 3D using simple, free-moving, non-directional detector systems.

The use of LiDAR-SLAM may be prohibitive in some environments such as feature-less underground tunnels or in areas where there are limitations on active sensors. The capabilities of other contextual data streams (e.g., RGB cameras, stereo depth cameras, GPS/INS, etc.) should be explored with new localization and mapping techniques being developed in the field of computer vision. For example, recent work [162] using deep neural networks has generated high-resolution 3D spatial meshes from single RGB images.

The experimental results presented in this work showed images in units of relative intensity and thus continued effort is required to produce quantitative images in units of absolute gamma-ray activity. Ultimately, quantitative source emission maps could be used to generate 3D dose maps for uses in, for example, emergency response and consequence management. Detailed experimental response characterizations will be necessary as well as the development of high-fidelity benchmarked simulations.

Finally, while 3D imaging was shown to perform in real-time for the measurements in this work, the reconstruction slows with the addition of more contextual and gamma-ray data. Mapping large scale, high-count rate environments over longer periods of time therefore requires a more scalable reconstruction approach to maintain real-time performance. This could include, for example, the implementation of a real-time probabilistic tri-state occupancy model and an adaptive non-uniform voxelization scheme.

Bibliography

[1] E. Caroli, J. B. Stephen, G. Di Cocco, L. Natalucci, and A. Spizzichino, "Coded Aperture Imaging in X- and Gamma-ray Astronomy," *Space Sci. Rev.*, vol. 45, pp. 349–403, 1987.

[2] H. Chang, S. Boggs, and Y. Chang, "The Nuclear Compton Telescope (NCT): Scientific Goals and Expected Sensitivity," *Adv. Space Res.*, vol. 40, no. 8, pp. 1281–1287, 2007.

[3] M. S. Bandstra *et al.*, "Detection and Imaging of the Crab Nebula with the Nuclear Compton Telescope," *Astrophys. J.*, vol. 738, no. 8, 2011.

[4] H. O. Anger, "Scintillation Camera with Multichannel Collimators," *J. Nucl. Med.*, vol. 5, pp. 515–531, 1964.

[5] R. Accorsi, "Design of Near-field Coded Aperture Cameras for High-Resolution Medical and Industrial Gamma-ray Imaging," PhD thesis, Massachusetts Institute of Technology, 2001.

[6] X. Zhang, Y. Lu, and T. Chan, "A Novel Sparsity Reconstruction Method from Poisson Data for 3D Bioluminescence Tomography," *J. Sci. Comput.*, vol. 50, no. 3, pp. 519–535, 2012.

[7] Z. Mu, L. W. Dobrucki, and Y. Liu, "SPECT Imaging of 2-D and 3-D Distributed Sources with Near-Field Coded Aperture Collimation: Computer Simulation and Real Data Validation," *J. Med. Biol. Eng.*, vol. 36, pp. 32–43, 2016.

[8] Y. Sato *et al.*, "Remote Radiation Imaging System using a Compact Gamma-Ray Imager Mounted on a Multicopter Drone," *J. Nucl. Sci. Technol.*, vol. 55, no. 1, pp. 90–96, 2018.

[9] K. Vetter, R. Barnowski, J. W. Cates, A. Haefner, T. H. Y. Joshi, R. Pavlovsky, and B. J. Quiter, "Advances in Nuclear Radiation Sensing: Enabling 3-D Gamma-Ray Vision," *Sensors*, vol. 19, no. 11, 2019.

[10] K. P. Ziock, C. J. Hailey, T. B. Gosnell, and J. H. Lupton, "A Gamma-Ray Imager for Arms Control," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 4, pp. 1046–1050, 1992.

[11] D. Xu and Z. He, "Gamma-ray Energy-Imaging Integrated Spectral Deconvolution," *Nucl. Instrum. Meth. A*, vol. 574, pp. 98–109, 2007.

[12] L. Mertz and N. O. Young, "Fresnel transformations of images," *in Proc. Int. Conf. Optical Instrum. Tech.*, p. 305, 1961.

[13] R. H. Dicke, "Scatter-Hole Cameras for X-rays and Gamma-rays," *Astrophys. J.*, vol. 153, pp. 101–106, 1968.

[14] J. G. Ables, "Fourier Transform Photography: A New Method for X-ray Astronomy," *Publ. Astronom. Soc. Australia*, vol. 1, no. 4, pp. 172–173, 1968.

[15] M. J. Cieślak, K. A. A. Gamage, and R. Glover, "Coded-Aperture Imaging Systems: Past, Present and Future Development - A Review," *Radiat. Meas.*, vol. 92, pp. 59–71, 2016.

[16] E. E. Fenimore and T. M. Cannon, "Coded Aperture Imaging with Uniformly Redundant Arrays," *Appl. Optics*, vol. 17, no. 3, pp. 337–347, 1978.

[17] K. P. Ziock, M. Cunningham, and L. Fabris, "Two-Sided Coded-Aperture Imaging Without a Detector Plane," *in Proc. IEEE Nucl. Sci. Symp.*, pp. 634–641, 2008.

[18] M. Galloway, A. Zoglauer, M. Amman, S. E. Boggs, and P. N. Luke, "Simulation and Detector Response for the High Efficiency Multimode Imager," *Nucl. Instrum. Meth. A*, vol. 652, pp. 641–645, 2011.

[19] A. Haefner, R. Barnowski, P. N. Luke, M. Amman, and K. Vetter, "Handheld Real-time Volumetric 3-D Gamma-ray Imaging," *Nucl. Instrum. Meth. A*, vol. 857, pp. 42–49, 2017.

[20] R. Arlt *et al.*, "Use of CdZnTe Detectors in Hand-Held and Portable Isotope Identifiers to Detect Illicit, Trafficking of Nuclear Material and Radioactive Sources," *in Proc. IEEE Nucl. Sci. Symp.*, 2000.

[21] J. Chin, D. K. Y. Yau, and N. S. V. Rao, "Efficient and Robust Localization of Multiple Radiation Sources in Complex Environments," *Int. Conf. Dis. Comput. Sys.*, pp. 780–789, 2011.

[22] R. T. Klann, S. C. De la Barrera, and R. B. Vilim, "Treatment of Shielding in Real-Time Source Tracking Software," *Nucl. Tech.*, vol. 175, pp. 301–313, 2011.

[23] K. N. Shokhirev, D. Konno, T. M. Schmit, V. Ziskin, and B. R. Cosofret, "Man-portable Radiation Detector based on Advanced Source Detection, Identification, and Localization Algorithms," *in Proc. Nucl. Sci. Symp. Med. Imaging Conf.*, 2015.

[24] F. E. Schneider and D. Wildermuth, "An Autonomous Unmanned Vehicle for CBRNE Reconnaissance," *in Proc. Int. Carpath. Control Conf.*, pp. 347–352, 2011.

[25] P. Jasiobedzki, H. Ng, M. Bondy, and C. H. McDiarmid, "C2SM: A Mobile System of Detecting and 3D Mapping of Chemical, Radiological, and Nuclear Contamination," *in Proc. SPIE*, vol. 7305, 2009.

[26] Y. Sanada and T. Torii, "Aerial Radiation Monitoring around the Fukushima Daiichi Nuclear Power Plant using an Unmanned Helicopter," *J. Environ. Radioactiv.*, vol. 139, no. 294-299, 2015.

[27] P. .G. Martin, O. D. Payton, J. S. Fardoulis, D. A. Richards, and T. B. Scott, "The Use of Unmanned Aerial Systems for the Mapping of Legacy Uranium Mines," *J. Environ. Radioactiv.*, vol. 143, no. 135-140, 2015.

[28] P. .G. Martin *et al.*, "3D Unmanned Aerial Vehicle Radiation Mapping for Assessing Contaminant Distribution and Mobility," *Int. J. Appl. Earth Obs.*, vol. 52, pp. 12–19, 2016.

[29] R. Pavlovsky *et al.*, "3-D Radiation Mapping in Real-Time with the Localization and Mapping Platform LAMP from Unmanned Aerial Systems and Man-Portable Configurations," *arXiv:1901.05038*, 2018.

[30] C. G. Wahl *et al.*, "The Polaris-H Imaging Spectrometer," *Nucl. Instrum. Meth. A*, vol. 784, pp. 377–381, 2015.

[31] *PHDS Co. – Gamma Ray Imaging Detectors*. [Online]. Available: `http://phdsco.com/`.

[32] L. Mihailescu, K. Vetter, and D. Chivers, "Standoff 3D Gamma-Ray Imaging," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 2, pp. 479–486, 2009.

[33] A. Kishimoto, J. Kataoka, T. Nishiyama, T. Taya, and S. Kabuki, "Demonstration of Three-Dimensional Imaging based on Handheld Compton Camera," *J. Instrum.*, vol. 10, 2015.

[34] H. Durrant-Whyte and T. Bailey, "Simultaneous Localization and Mapping: Part I," *IEEE Robot. Autom.*, vol. 13, no. 2, 2006.

[35] ——, "Simultaneous Localization and Mapping: Part II," *IEEE Robot. Autom.*, vol. 13, no. 3, 2006.

[36] R. Barnowski, A. Haefner, L. Mihailescu, and K. Vetter, "Scene Data Fusion - Real-time Standoff Volumetric Gamma-ray Imaging," *Nucl. Instrum. Methods A*, vol. 800, pp. 65–69, 2015.

[37] M. B. Pellazar, "Vehicle Route Planning with Constraints using Genetic Algorithms," *in Proc. Nat. Aerospace Elec. Conf.*, 1994.

[38] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *J. R. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.

[39] L. Parra and H. H. Barrett, "List-Mode Likelihood: EM Algorithm and Image Quality Estimation Demonstrated on 2-D PET," *IEEE Trans. Med. Imaging*, vol. 17, no. 2, pp. 228–235, 1998.

[40] R. Barnowski, "Development and Evaluation of Real-Time Volumetric Compton Gamma-Ray Imaging," PhD thesis, University of California, Berkeley, 2016.

[41] D. Hellfeld, P. Barton, D. Gunter, L. Mihailescu, and K. Vetter, "A Spherical Active Coded Aperutre for 4π Gamma-ray Imaging," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 11, pp. 2837–2842, 2017.

[42] D. Hellfeld, T. H. Y. Joshi, M. S. Bandstra, R. J. Cooper, B. J. Quiter, and K. Vetter, "Gamma-Ray Point-Source Localization and Sparse Image Reconstruction using Poisson Likelihood," *IEEE Trans. Nucl. Sci. (in press)*, 2019.

[43] D. Hellfeld, P. Barton, D. Gunter, A. Haefner, L. Mihailescu, and K. Vetter, "Real-time Free-moving Active Coded Mask 3D Gamma-Ray Imaging," *Submitted to IEEE Trans. Nucl. Sci. (under review)*, 2019.

[44] G. F. Knoll, *Radiation Detection and Measurement*, 4^{th} edition. John Wiley & Sons, 2010.

[45] K. S. Krane, *Introductory Nuclear Physics*. John Wiley & Sons, Inc., 1988.

[46] J. Lilley, *Nuclear Physics: Principles and Applications*. John Wiley & Sons, Ltd., 2001.

[47] J. E. Turner, *Atoms, Radiation, and Radiation Protection*. WILEY-VCH, 2007.

[48] G. Gilmore, *Practical Gamma-ray Spectroscopy*. John Wiley & Sons, Ltd., 2008.

[49] L. C. Maximon and H. A. Gimm, "Pair Production in the Field of Atomic Electrons," *Phys. Rev. A*, vol. 23, no. 1, pp. 172–185, 1981.

[50] H. Spieler, *Semiconductor Detector Systems*. Oxford University Press, 2005.

[51] B. F. Phlips, W. N. Johnson, R. A. Kroeger, J. D. Kurfess, G. Phillips, E. A. Wulf, and P. N. Luke, "Development of Germanium Strip Detectors for Environmental Remediation," *IEEE Trans. Nucl. Sci.*, vol. 49, no. 2, pp. 597–600, 2002.

[52] L. Mihailescu, K. M. Vetter, M. T. Burks, E. L. Hull, and W. W. Craig, "SPEIR: A Ge Compton Camera," *Nucl. Instrum. Meth. A*, vol. 570, pp. 89–100, 2006.

[53] P. N. Luke, "Single-polarity Charge Sensing in Ionization Detectors using Coplanar Electrodes," *Appl. Phys. Lett.*, vol. 65, no. 22, pp. 2884–1886, 1994.

[54] ——, "Unipolar Charge Sensing with Coplanar Electrodes - Application to Semiconductor Detectors," *IEEE Trans. Nucl. Sci.*, vol. 42, no. 4, pp. 207–213, 1995.

[55] ——, "Electrode configuration and energy resolution in gamma-ray detectors," *Nucl. Instrum. Meth. A*, vol. 380, pp. 232–237, 1996.

[56] P. N. Luke, M. Amman, T. H. Prettyman, P. A. Russo, and D. A. Close, "Electrode design for coplanar-grid detectors," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 713–720, 1997.

[57] Z. He, G. F. Knoll, D. K. Wehe, and J. Miyamoto, "Position-sensitive Single Carrier CdZnTe Detectors," *Nucl. Instrum. Meth. A*, vol. 388, pp. 180–185, 1997.

[58] F. P. Doty *et al.*, "Pixellated CdZnTe Detector Arrays," *Nucl. Instrum. Meth. A*, vol. 353, pp. 356–360, 1994.

[59] F. Zhang, Z. He, D. Xu, G. F. Knoll, D. K. Wehe, and J. E. Berry, "Improved Resolution for 3-D Position Sensitive CdZnTe Spectrometers," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2427–2431, 2004.

[60] Z. He, W. Li, G. F. Knoll, D. K. Wehe, and C. M. Stahle, "Measurement of Material Uniformity using 3-D Position Sensitive CdZnTe Gamma-ray Spectrometers," *Nucl. Instrum. Meth. A*, vol. 441, pp. 459–467, 2000.

[61] N. F. Brejnholt *et al.*, "Demonstration of Multilayer Reflective Optics at Photon Energies Above 0.6 MeV," *Optics Express*, vol. 22, no. 13, pp. 15 364–15 369, 2014.

[62] P. von Ballmoos *et al.*, "CLAIRE: First Light for a Gamma-ray Lens," *Exp. Astron.*, vol. 20, no. 253–267, 2005.

[63] J. E. Grindlay and J. Hong, "Optimizing Wide-Field Coded Aperture Imaging: Radial Mask Holes and Scanning," *in Proc. SPIE*, vol. 402, 2004.

[64] J. Hong *et al.*, "Laboratory Coded-Apeture Imaging Experiments: Radial Hole Coded Masks and Depth-Sensitive CZT Detectors," *Proc. SPIE 5540, Hard X-Ray and Gamma-Ray Detector Physics VI*, vol. 63, 2004.

[65] K. Amgarou *et al.*, "A Comprehensive Experimental Characterization of the iPIX Gamma Imager," *J. Instrum.*, vol. 11, 2016.

[66] C. E. Lehner, Z. He, and F. Zhang, "4π Compton Imaging Using a 3-D Position-Sensitive CdZnTe Detector via Weighted List-Mode Maximum Likelihood," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 4, pp. 1618–1624, 2004.

[67] A. Haefner, D. Gunter, R. Barnowski, and K. Vetter, "A Filtered Back-Projection Algorithm for 4π Comton Camera Data," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 4, pp. 1911–1917, 2015.

[68] A. C. Raffo-Caiado *et al.*, "Combining Measurements with Three-Dimensional Laser Scanning System and Coded Aperture Gamma-Ray Imaging System for International Safeguards Applications," *IAEA Symp. Int. Safeguards*, vol. 41, no. 50, 2010.

[69] F. Enres, J. Hess, N. Engelhard, J. Sturm, D. Cremers, and W. Burgard, "An Evaluation of the RGB-D SLAM System," *IEEE Int. Conf. Robot. Autom.*, 2012.

[70] D. Kim, H. Woo, Y. Ji, Y. Tamura, A. Yamashita, and H. Asama, "3D Radiation Imaging Using Mobile Robot Equipped with Radiation Detector," *in Proc. IEEE/SICE Int. Symp. Sys. Integr.*, 2017.

[71] K. Vetter, R. Barnowski, A. Haefner, T. H. Y. Joshi, R. Pavlovsky, and B. J. Quiter, "Gamma-Ray Imaging for Nuclear Security and Safety: Towards 3-D Gamma-Ray Vision," *Nucl. Instrum. Meth. A*, vol. 878, pp. 159–168, 2018.

[72] *Velodyne LiDAR - Puck LITE*. [Online]. Available: `https://velodynelidar.com/vlp-16-lite.html`.

[73] *VectorNav Embedded Navigation Solutions - VN-200 GPS-Aided INS*. [Online]. Available: `https://www.vectornav.com/products/vn-200`.

[74] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-Time Loop Closure in 2D LIDAR SLAM," *in Proc. IEEE Int. Conf. Robot. Autom.*, pp. 1271–1278, 2016.

[75] M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger, "Real-time 3D Reconstruction at Scale using Voxel Hashing," *ACM Trans. Graphics*, vol. 32, no. 6, 2013.

[76] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An Efficient Probabilistic 3D Mapping Framework based on Octrees," *Autonom. Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[77] *Agisoft Metashape Standard (Version 1.5.2) (Software). (2019)*. [Online]. Available: `http://www.agisoft.com/downloads/installer/`.

[78] M. Jancosek and T. Pajdla, "Multi-view Reconstruction Preserving Weakly-Supported Surfaces," *in Proc. Conf. Computer Vision and Pattern Recognition*, 2011.

[79] P. Moulon, P. Monasse, and R. Marlet, "Adaptive Structure from Motion with a Contrario Model Estimation," *in Proc. Asian Conf. Computer Vision*, pp. 257–270, 2012.

[80] J. L. Schönberger and J. Frahm, "Structure-from-Motion Revisited," *in Proc. Conf. Computer Vision and Pattern Recognition*, 2016.

[81] J. L. Schönberger, E. Zheng, M. Pollefeys, and J. Frahm, "Pixelwise View Selection for Unstructured Multi-View Stereo," *in Proc. European Conf. Computer Vision*, 2016.

[82] K. J. Bilton *et al.*, "Non-negative Matrix Factorization of Gamma-ray Spectra for Background Modeling, Detection, and Source Identification," *IEEE Trans. Nucl. Sci.*, vol. 68, no. 5, pp. 827–837, 2019.

[83] J. T. Bushberg, J. A. Seibert, J. E. M. Leidholdt, and J. M. Boone, *The Essential Physics of Medical Imaging*, 3rd edition. Lippincot Williams & Wilkins, 2012.

[84] M. Conti and M. E. Casey, "Estimating the Optimal Iteration Number in Iterative Reconstruction: A Statistical Approach," *in Proc. IEEE Nucl. Sci. Symp.*, 2007.

[85] C. Pafilis *et al.*, "A Methodology for the Estimation of the Optimal Iteration in MLEM-based Image Reconstruction in PET," *in Proc. Int. Workshop Biomed. Eng.*, 2001.

[86] P. J. Green, "Bayseian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm," *IEEE Trans. Med. Imaging*, vol. 9, no. 1, pp. 84–93, 1990.

[87] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, "Total Variation Regulated EM Algorithm," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2202–2210, 1999.

[88] S. van der Walt *et al.*, "scikit-image: Image Processing in Python," *PeerJ 2:e453*, 2014.

[89] E. J. Candès, J. Romberg, and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[90] A. Haefner, "Compton Image Reconstruction Algorithms and Demonstration Across Multiple Devices: From the Lab to the Field," PhD thesis, University of California, Berkeley, 2014.

[91] A. K. Fletcher, K. Ramchandran, and V. K. Goyal, "Wavelet Denoising by Recursive Cycle Spinning," *in Proc. Int. Conf. Image Process.*, 2002.

[92] D. J. Lingenfelter, J. A. Fessler, and Z. He, "Sparsity Regularization for Image Reconstruction with Poisson Data," *in Proc. SPIE*, vol. 7246, 2009.

[93] K. Lange, M. Bahn, and R. Little, "A Theoretical Study of Some Maximum Likelihood Algorithms for Emission and Transmission Tomography," *IEEE Trans. Med. Imaging*, vol. 6, no. 2, pp. 106–114, 1987.

[94] C. Byrne, "Likelihood Maxmimuzation for List-Mode Emission Tomographic Image Reconstruction," *IEEE Trans. Med. Imaging*, vol. 20, no. 10, pp. 1084–1092, 2001.

[95] J. E. Stone, D. Gohara, and G. Shi, "OpenCL: A Parallel Programming Standard for Heterogeneous Computing Systems," *Comput. Sci. Eng.*, vol. 12, no. 3, pp. 66–73, 2010.

[96] A. Klöckner, N. Pinto, Y. Lee, B. Catanzaro, P. Ivanov, and A. Fasih, "PyCUDA and PyOpenCL: A Scripting-Based Approach to GPU Run-Time Code Generation," *Parallel Comput.*, vol. 38, no. 3, pp. 157–174, 2012.

[97] J. Nickolls, I. Buck, M. Garland, and K. Skadron, "Scalable Parallel Programming with CUDA," *ACM Queue - GPU Comput.*, vol. 6, no. 2, pp. 40–53, 2008.

[98] H. M. Hudson and R. S. Larkin, "Accelerated Image Reconstruction Using Ordered Subset of Projection Data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[99] A. Rahmim *et al.*, "Statistical List-mode Image Reconstruction for the High Resolution Research Tomograph," *Phys. Med. Biol.*, vol. 49, pp. 4239–4258, 2004.

[100] G. Pratx, G. Chinn, F. Habte, P. Olcott, and C. Levin, "Fully 3-D List-Mode OSEM Accelerated by Graphics Processing Units," *in Proc. IEEE Nucl. Sci. Symp.*, 2006.

[101] S. Ahn and J. A. Fessler, "Globally Convergent Image Reconstruction for Emission Tomography Using Relaxed Ordered Subsets Algorithms," *IEEE Trans. Med. Imaging*, vol. 22, no. 5, pp. 613–626, 2003.

[102] P. Khurd, I. Hsiao, A. Rangarajan, and G. Gindi, "A Globally Convergent Regularized Ordered-Subset EM Algorithm for List-Mode Reconstruction," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 3, pp. 719–725, 2004.

[103] G. De Geronimo, G. Carini, W. S. Murray, and P. O'Connor, "Front-End ASIC for Co-Planar Grid Sensors," *IEEE Trans. Nucl. Sci.*, vol. 52, no. 5, pp. 2003–2008, 2005.

[104] M. Turqueti, R. A. Rivera, A. Prosser, J. Andersen, and J. Chramowicz, "CAPTAN: A Hardware Architecture for Integrated Data Acquisition, Control, and Analysis for Detector Development," *In Proc: IEEE Nucl. Sci. Symp.*, 2008.

[105] *Intel NUC7i7BNH*. [Online]. Available: `https://www.intel.com/content/www/us/en/products/boards-kits/nuc/kits/nuc7i7bnh.html?wapkw=nuc7i7bnh`.

[106] *FLIR - Flea3 1.3 MP Color USB3 Vision*. [Online]. Available: `https://www.ptgrey.com/flea3-13-mp-color-usb3-vision-e2v-ev76c560-camera`.

[107] *Kowa - LM4NCL Lens*. [Online]. Available: `https://lenses.kowa-usa.com/ncl-series/488-lm4ncl.html`.

[108] S. Agostinelli *et al.*, "Geant4 - A Simulation Toolkit," *Nucl. Instrum. Meth. A*, vol. 506, pp. 250–303, 2003.

[109] K. M. Górski *et al.*, "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere," *Astrophys. J.*, vol. 662, no. 2, pp. 759–771, 2005.

[110] E. E. Fenimore, "Coded Aperture Imaging Predicted Performance of Uniformly Redundant Arrays," *Appl. Optics*, vol. 17, no. 22, pp. 3562–3570, 1978.

[111] A. Busboom, H. Elders-Boll, and H. D. Schotten, "Combinatorial Design of Near-optimum Masks for Coded Aperture Imaging," *in Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, pp. 2817–2820, 1997.

[112] A. Zoglauer, M. Galloway, M. Amman, S. E. Boggs, and P. N. Luke, "HEMI's Coded Mask: Pattern Selection, Imaging Approaches, and Performance," *in Proc. IEEE Symp. Radiat. Meas. App.*, 2012.

[113] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science, New Series*, vol. 220, no. 4598, pp. 671–680, 1983.

[114] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, "Equation of State Calculation by Fast Computing Machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, 1953.

[115] W. K. Hastings, "Monte Carlo Sampling Methods using Markov Chains and their Applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[116] G. Dueck, "New Optimization Heuristics - The Great Deluge Algorithm and the Record-to-Record Travel," *J. Comp. Phys.*, vol. 104, no. 1, pp. 86–92, 1993.

[117] W. Wang, J. M. Jaworski, W. R. Kaye, and Z. He, "Applications of the Energy-Imaging Integrated Deconvolution Algorithm for Source Characterization," *In Proc. IEEE Nucl. Sci. Symp.*, 2009.

[118] B. Yu, S. Brown, Z. He, and J. Cheng, "Application of Chambolle-Pock Algorithm on Penalized Gamma-ray Energy-Imaging Integrated Deconvolution," *in Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2015.

[119] D. Reilly, N. Ensslin, H. Smith, and S. Keiner, *Passive Nondestructive Assay of Nuclear Materials (PANDA)*. United States Nuclear Regulatory Comission, 1991.

[120] M. Streicher, S. Brown, Y. Zhu, D. Goodman, and Z. He, "Special Nuclear Material Characterization Using Digital 3-D Position Sensitive CdZnTe Detectors and High Purity Germanium Spectrometers," *IEEE Trans. Nucl. Sci.*, vol. 63, no. 5, pp. 2649–2656, 2016.

[121] M. A. Alnafea, K. Wells, M. Guy, and N. M. Spyrou, "Near Field Corrections for Coded Aperture Imaging in Scintimmammography," *in Proc. IEEE Nucl. Sci. Symp.*, 2006.

[122] X. Hubert, D. Chambellan, S. Legoupil, J. Deverre, and N. Paragios, "Direct 3D-Reconstruction from a Single 2D-image Obtained Through a Coded-Aperture," *in Proc. IEEE Nucl. Sci. Symp.*, 2007.

[123] Z. Mu, W. L. Dobrucki, X. Hu, and Y. Liu, "Recent Progress on SPECT Imaging with Near-field Coded Aperture Collimation: A Small Animal Study," *in Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2010.

[124] K. P. Ziock, J. B. Braverman, L. Fabris, and M. J. Harrison, "The Other End of the Scale: Coded Apertures in the Near Field for High Resolution 3D Gamma Event Localization in Bulk Scintillators," *in Proc. SPIE*, 2012.

[125] D. Shreiner, G. Sellers, J. Kessenich, and B. Licea-Kane, *OpenGL Programming Guide*, 8th edition. Addison-Wesley, 2013.

[126] M. Kyto, M. Nuutinen, and P. Oittinen, "Method for Measuring Stereo Camera Depth Accuracy based on Stereoscopic Vision," *in Proc. SPIE*, 2011.

[127] K. Khoshelham and S. O. Elberink, "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.

[128] M. Amman, P. N. Luke, and J. S. Lee, "CdZnTe Material Uniformity and Coplanar-Grid Gamma-Ray Detector Performance," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 3, pp. 760–767, 2000.

[129] K. Zahraman *et al.*, "Study of the Thickness of the Dead Layer Below Electrodes, Deposited by Electroless Technique, in CdTe Nuclear Detectors," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 1, pp. 378–382, 2006.

[130] N. S. V. Rao *et al.*, "Identification of Low-Level Point Radiation Sources Using a Sensor Network," *in Proc. Int Conf. Inf. Process. Sens. Networks*, pp. 493–504, 2008.

[131] B. Deb, J. A. F. Ross, A. Ivan, and M. J. Hartman, "Radioactive Source Estimation Using a System of Directional and Non-directional Detectors," *IEEE Trans. Nucl. Sci.*, vol. 58, no. 6, pp. 3281–3290, 2011.

[132] B. Deb, "Iterative Estimation of Location and Trajectory of Radioactive Sources with a Networked System of Detectors," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 1315–1326, 2013.

[133] U. Bhattacharyya and C. Baum, "Estimating the Location of a Nuclear Source in a Three-Dimensional Environment Using a Two-stage Adaptive Algorithm," *in Proc. of IEEE Comput. Comm. Workshop Conf.*, 2018.

[134] E. A. Miller *et al.*, "Adaptively Reevaluated Bayesian Localization (ARBL): A Novel Technique for Radiological Source Localization," *Nucl. Instrum. Meth. A*, vol. 785, pp. 332–338, 2015.

[135] B. Ristic, M. Morelande, and A. Gunatilaka, "Information Driven Search for Point Sources of Gamma Radiation," *Signal Process.*, vol. 90, no. 4, pp. 1225–1239, 2010.

[136] J. Chin, D. K. Yau, N. S. V. Rao, Y. Yang, C. Y. T. Ma, and M. Shankar, "Accurate Localization of Low-level Radioactive Source Under Noise and Measurement Errors," *in Proc. ACM Conf. Embed. Network Sens. Sys.*, pp. 183–196, 2008.

[137] R. Vilim and R. Klann, "RadTrac: A System for Detecting, Localizing, and Tracking Radioactive Sources in Real Time," *Nucl. Tech.*, vol. 168, no. 1, pp. 61–73, 2009.

[138] M. K. Sharma, A. B. Alajo, and H. K. Lee, "Three-dimensional Localization of Low Activity Gamma-ray Sources in Real-time Scenarios," *Nucl. Instrum. Meth. A*, vol. 813, no. 21, pp. 132–138, 2016.

[139] G. Cordone, R. R. Brooks, S. Sen, N. S. V. Rao, C. Q. Wu, M. L. Berry, and K. M. Grieme, "Improved Multi-Resolution Method for MLE-based Localization of Radiation Sources," *in Proc. Int. Conf. Inf. Fusion*, 2017.

[140] J. C. Curtis *et al.*, "Simulation and Validation of the Mobile Urban Radiation Search (MURS) Gamma-ray Detector Response," *Nucl. Instrum. Meth. A*, 2018.

[141] T. H. Y. Joshi *et al.*, "Measurement of the Energy Dependent Angular Response of the ARES Detector System and Application to Aerial Imaging," *IEEE Trans. Nucl. Sci.*, vol. 64, no. 7, pp. 1754–1760, 2017.

[142] Z. Harmany, R. F. Marcia, and R. M. Willet, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms - Theory and Practice," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1084–1096, 2012.

[143] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Ann. Math. Stat.*, vol. 9, no. 1, pp. 60–62, 1938.

[144] M. J. D. Powell, "The BOBYQA Algorithm for Bound Constrained Optimization Without Derivatives," *Department of Applied Mathematics and Theoretical Physics, Cambridge University Technical Report NA2009/06*, 2009.

[145] S. G. Johnson, "The NLopt Nonlinear-Optimization Package," 2008. [Online]. Available: `http://ab-initio.mit.edu/nlopt`.

[146] F. Biscani *et al.*, "Esa/Pagmo2: Pagmo 2.10," 2019. [Online]. Available: `https://doi.org/10.5281/zenodo.2529931`.

[147] N. Boyd, G. Schiebinger, and B. Recht, "The Alternating Descent Conditional Gradient Method for Sparse Inverse Problems," *SIAM J. Optim.*, vol. 27, no. 2, pp. 616–639, 2017.

[148] G. Schwarz, "Estimating the Dimension of a Model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1974.

[149] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[150] R. Rosson, J. Lahr, and B. Kahn, "Radiation Background in a $LaBr_3$(Ce) Gamma-Ray Scintillation Detector," *Health Phys.*, vol. 101, no. 6, pp. 703–708, 2011.

[151] T. J. Aucott, M. S. Bandstra, V. Negut, D. H. Chivers, R. J. Cooper, and K. Vetter, "Routine Surveys for Gamma-Ray Background Characterization," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 2, pp. 1147–1150, 2013.

[152] R. D. Penny, "Improved Radiological/Nuclear Source Localization in Variable NORM Background: An MLEM Approach with Segmentation Data," *Nucl. Instrum. Meth. A*, vol. 784, pp. 319–325, 2015.

[153] E. Bai, A. Heifetz, P. Raptis, S. Dasgupta, and R. Mudumbai, "Maximum Likelihood Localization of Radioactive Sources Against a Highly Fluctuating Background," *IEEE Trans. Nucl. Sci.*, vol. 62, no. 6, pp. 3274–3281, 2015.

[154] K. Henderson, K. Stadnikia, A. Martin, A. Enqvist, and S. Koppal, "Tracking Radioactive Sources through Sensor Fusion of Omnidirectional LIDAR and Isotropic Rad-detectors," *in Proc. Int. Conf. 3D Vision*, 2017.

[155] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "PointPillars: Fast Encoders for Object Detection from Point Clouds," *arXiv:1812.05784v2*, 2018.

[156] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple Online and Realtime Tracking," *arXiv:1602.00763v2*, 2017.

[157] U. Shirwadkar, R. Hawrami, J. Glodo, E. V. D. van Loef, and K. S. Shah, "Novel Scintillation Material $Cs_2LiLaBr_{6-x}Cl_x$:Ce for Gamma-ray and Neutron Spectroscopy," *in Proc. IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2012.

[158] R. Hawrami, L. S. Pandian, E. Ariesanti, J. Glodo, J. Finkelstein, J. Tower, and K. Shah, "$Cs_2LiLa(Br, Cl)_6$ Crystals for Nuclear Security Applications," *IEEE Trans. Nucl. Sci.*, vol. 63, no. 2, pp. 509–512, 2016.

[159] J. Glodo, W. M. Higgins, E. V. D. van Loef, and K. S. Shah, "Cs_2LiYCl_6:Ce Scintillator for Nuclear Monitoring Applications," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 3, pp. 1257–1261, 2009.

[160] A. Bieberle, T. Schäfer, M. Neumann, and U. Hampel, "Validation of High-Resolution Gamma-ray Computed Tomography for Quantitative Gas Holdup Measurements in Centrifugal Pumps," *Meas. Sci. Tech.*, vol. 26, no. 9, 2015.

[161] T. J. Aucott, A. D. Brand, and D. P. DiPrete, "Gamma-ray Imaging and Holdup Assays of 235-F PuFF cells 1 & 2," *SRNL Technical Report*, 2017.

[162] N. Wang, Y. Zhang, Z. Li, Y. Fu, W. Liu, and Y. Jiang, "Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images," *arXiv:1804.01654v2*, 2018.